Constantine Stephanidis
Margherita Antona
Stavroula Ntoa
Gavriel Salvendy (Eds.)

Communications in Computer and Information Science 2529

HCI International 2025 Posters

27th International Conference on
Human-Computer Interaction, HCII 2025
Gothenburg, Sweden, June 22–27, 2025
Proceedings, Part VIII

Part 8

Springer

Communications in Computer and Information Science 2529

Series Editors

Gang Li, *School of Information Technology, Deakin University, Burwood, VIC, Australia*
Joaquim Filipe, *Polytechnic Institute of Setúbal, Setúbal, Portugal*
Zhiwei Xu, *Chinese Academy of Sciences, Beijing, China*

Rationale
The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics
The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors
Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (http://link.springer.com) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at http://link.springer.com/bookseries/7899. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process
The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing
CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start
To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com

Constantine Stephanidis · Margherita Antona ·
Stavroula Ntoa · Gavriel Salvendy
Editors

HCI International 2025 Posters

27th International Conference on
Human-Computer Interaction, HCII 2025
Gothenburg, Sweden, June 22–27, 2025
Proceedings, Part VIII

Springer

Editors
Constantine Stephanidis
University of Crete and Foundation for Research and Technology - Hellas (FORTH)
Heraklion, Crete, Greece

Margherita Antona
Foundation for Research and Technology - Hellas (FORTH)
Heraklion, Crete, Greece

Stavroula Ntoa
Foundation for Research and Technology - Hellas (FORTH)
Heraklion, Crete, Greece

Gavriel Salvendy
University of Central Florida
Orlando, FL, USA

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-94170-2 ISBN 978-3-031-94171-9 (eBook)
https://doi.org/10.1007/978-3-031-94171-9

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Foreword

The HCI International (HCII) conference was founded in 1984 by Gavriel Salvendy (Purdue University, USA, Tsinghua University, P.R. China, and University of Central Florida, USA) and the first event of the series, "1st USA-Japan Conference on Human-Computer Interaction", was held in Honolulu, Hawaii, USA, 18–20 August. Since then, HCI International is held jointly with several Thematic Areas and Affiliated Conferences, with each one under the auspices of a distinguished international Program Board and under one management and one registration. Twenty-seven HCI International Conferences have been organized so far (every two years until 2013, and annually thereafter).

Last year, we celebrated 40 years since the establishment of the HCII conference, which has been a hub for presenting groundbreaking research and novel ideas and collaboration for people from all over the world. Over the years, this conference has served as a platform for scholars, researchers, industry experts, and students to exchange ideas, connect, and address challenges in the ever-evolving HCI field. The conference has evolved itself, adapting to new technologies and emerging trends, while staying committed to its core mission of advancing knowledge and driving change.

The 27th International Conference on Human-Computer Interaction, HCI International 2025 (HCII 2025), was held as an 'on-site' conference at the Gothia Towers Hotel and Swedish Exhibition & Congress Centre, in Gothenburg, Sweden, on June 22–27, 2025, with the additional option for 'on-line' participation. It incorporated the 21 thematic areas and affiliated conferences listed below.

A total of 7972 individuals from academia, research institutes, industry, and government agencies from 92 countries submitted contributions. 1430 papers and 355 posters (as short research papers) are included in the volumes of the proceedings published just before the start of the conference, and which are listed below. The contributions thoroughly cover the entire field of human-computer interaction, highlight the evolving role of computers in diverse contexts, and demonstrate how HCI research is shaping and improving user experiences across a wide range of domains, influencing technological progress and its effective integration into various sectors.

The HCII conference also offers the option of presenting 'Late Breaking Work', both for papers and posters, with the corresponding proceedings volumes published after the conference. Full papers are included in the 'HCII 2025 - Late Breaking Papers' volumes of the proceedings published in the Springer LNCS series, while 'Poster Extended Abstracts' are included as short research papers in the 'HCII 2025 - Late Breaking Posters' volumes published in the Springer CCIS series.

I would like to thank the Program Board Chairs and the members of the Program Boards of all thematic areas and affiliated conferences for their contribution towards the high scientific quality and overall success of the HCI International 2025 conference. Their manifold support including paper reviews (via a single-blind review process, with a minimum of two reviews per submission), session organization, and their willingness to act as goodwill ambassadors for the conference is most highly appreciated.

This conference would not have been possible without the continuous and unwavering support and advice of Gavriel Salvendy, founder, General Chair Emeritus, and Scientific Advisor. For his outstanding efforts, I would like to express my sincere appreciation to Abbas Moallem, Communications Chair and Editor of HCI International News.

June 2025 Constantine Stephanidis

HCI International 2025 Thematic Areas and Affiliated Conferences

- HCI: Human-Computer Interaction Thematic Area
- HIMI: Human Interface and the Management of Information Thematic Area
- EPCE: 22nd International Conference on Engineering Psychology and Cognitive Ergonomics
- AC: 19th International Conference on Augmented Cognition
- UAHCI: 19th International Conference on Universal Access in Human-Computer Interaction
- CCD: 17th International Conference on Cross-Cultural Design
- SCSM: 17th International Conference on Social Computing and Social Media
- VAMR: 17th International Conference on Virtual, Augmented and Mixed Reality
- DHM: 16th International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management
- DUXU: 14th International Conference on Design, User Experience, and Usability
- C&C: 13th International Conference on Culture and Computing
- DAPI: 13th International Conference on Distributed, Ambient and Pervasive Interactions
- HCIBGO: 12th International Conference on HCI in Business, Government and Organizations
- LCT: 12th International Conference on Learning and Collaboration Technologies
- ITAP: 11th International Conference on Human Aspects of IT for the Aged Population
- AIS: 7th International Conference on Adaptive Instructional Systems
- HCI-CPT: 7th International Conference on HCI for Cybersecurity, Privacy and Trust
- HCI-Games: 7th International Conference on HCI in Games
- MobiTAS: 7th International Conference on HCI in Mobility, Transport and Automotive Systems
- AI-HCI: 6th International Conference on Artificial Intelligence in HCI
- MOBILE: 6th International Conference on Human-Centered Design, Operation and Evaluation of Mobile Communications

List of Conference Proceedings Volumes Appearing Before the Conference

1. LNCS 15766, Human-Computer Interaction - Part I, edited by Masaaki Kurosu and Ayako Hashizume
2. LNCS 15767, Human-Computer Interaction - Part II, edited by Masaaki Kurosu and Ayako Hashizume
3. LNCS 15768, Human-Computer Interaction - Part III, edited by Masaaki Kurosu and Ayako Hashizume
4. LNCS 15769, Human-Computer Interaction - Part IV, edited by Masaaki Kurosu and Ayako Hashizume
5. LNCS 15770, Human-Computer Interaction - Part V, edited by Masaaki Kurosu and Ayako Hashizume
6. LNCS 15771, Human-Computer Interaction - Part VI, edited by Masaaki Kurosu and Ayako Hashizume
7. LNCS 15772, Human-Computer Interaction - Part VII, edited by Masaaki Kurosu and Ayako Hashizume
8. LNCS 15773, Human Interface and the Management of Information: Part I, edited by Hirohiko Mori and Yumi Asahi
9. LNCS 15774, Human Interface and the Management of Information: Part II, edited by Hirohiko Mori and Yumi Asahi
10. LNCS 15775, Human Interface and the Management of Information: Part III, edited by Hirohiko Mori and Yumi Asahi
11. LNAI 15776, Engineering Psychology and Cognitive Ergonomics: Part I, edited by Don Harris and Wen-Chin Li
12. LNAI 15777, Engineering Psychology and Cognitive Ergonomics: Part II, edited by Don Harris and Wen-Chin Li
13. LNAI 15778, Augmented Cognition: Part I, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
14. LNAI 15779, Augmented Cognition: Part II, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
15. LNCS 15780, Universal Access in Human-Computer Interaction: Part I, edited by Margherita Antona and Constantine Stephanidis
16. LNCS 15781, Universal Access in Human-Computer Interaction: Part II, edited by Margherita Antona and Constantine Stephanidis
17. LNCS 15782, Cross-Cultural Design: Part I, edited by Pei-Luen Patrick Rau
18. LNCS 15783, Cross-Cultural Design: Part II, edited by Pei-Luen Patrick Rau
19. LNCS 15784, Cross-Cultural Design: Part III, edited by Pei-Luen Patrick Rau
20. LNCS 15785, Cross-Cultural Design: Part IV, edited by Pei-Luen Patrick Rau

21. LNCS 15786, Social Computing and Social Media: Part I, edited by Adela Coman and Simona Vasilache
22. LNCS 15787, Social Computing and Social Media: Part II, edited by Adela Coman and Simona Vasilache
23. LNCS 15788, Virtual, Augmented and Mixed Reality: Part I, edited by Jessie Y. C. Chen and Gino Fragomeni
24. LNCS 15789, Virtual, Augmented and Mixed Reality: Part II, edited by Jessie Y. C. Chen and Gino Fragomeni
25. LNCS 15790, Virtual, Augmented and Mixed Reality: Part III, edited by Jessie Y. C. Chen and Gino Fragomeni
26. LNCS 15791, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Part I, edited by Vincent G. Duffy
27. LNCS 15792, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Part II, edited by Vincent G. Duffy
28. LNCS 15793, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Part III, edited by Vincent G. Duffy
29. LNCS 15794, Design, User Experience, and Usability: Part I, edited by Martin Schrepp
30. LNCS 15795, Design, User Experience, and Usability: Part II, edited by Martin Schrepp
31. LNCS 15796, Design, User Experience, and Usability: Part III, edited by Martin Schrepp
32. LNCS 15797, Design, User Experience, and Usability: Part IV, edited by Martin Schrepp
33. LNCS 15798, Design, User Experience, and Usability: Part V, edited by Martin Schrepp
34. LNCS 15799, Design, User Experience, and Usability: Part VI, edited by Martin Schrepp
35. LNCS 15800, Culture and Computing: Part I, edited by Matthias Rauterberg
36. LNCS 15801, Culture and Computing: Part II, edited by Matthias Rauterberg
37. LNCS 15802, Distributed, Ambient and Pervasive Interactions: Part I, edited by Norbert A. Streitz and Shin'ichi Konomi
38. LNCS 15803, Distributed, Ambient and Pervasive Interactions: Part II, edited by Norbert A. Streitz and Shin'ichi Konomi
39. LNCS 15804, HCI in Business, Government and Organizations: Part I, edited by Fiona Fui-Hoon Nah and Keng Leng Siau
40. LNCS 15805, HCI in Business, Government and Organizations: Part II, edited by Fiona Fui-Hoon Nah and Keng Leng Siau
41. LNCS 15806, Learning and Collaboration Technologies: Part I, edited by Brian K. Smith and Marcela Borge
42. LNCS 15807, Learning and Collaboration Technologies: Part II, edited by Brian K. Smith and Marcela Borge
43. LNCS 15808, Learning and Collaboration Technologies: Part III, edited by Brian K. Smith and Marcela Borge

44. LNCS 15809, Human Aspects of IT for the Aged Population: Part I, edited by Qin Gao and Jia Zhou
45. LNCS 15810, Human Aspects of IT for the Aged Population: Part II, edited by Qin Gao and Jia Zhou
46. LNCS 15811, Human Aspects of IT for the Aged Population: Part III, edited by Qin Gao and Jia Zhou
47. LNCS 15812, Adaptive Instructional System: Part I, edited by Robert A. Sottilare and Jessica Schwarz
48. LNCS 15813, Adaptive Instructional System: Part II, edited by Robert A. Sottilare and Jessica Schwarz
49. LNCS 15814, HCI for Cybersecurity, Privacy and Trust: Part I, edited by Abbas Moallem
50. LNCS 15815, HCI for Cybersecurity, Privacy and Trust: Part II, edited by Abbas Moallem
51. LNCS 15816, HCI in Games, edited by Xiaowen Fang
52. LNCS 15817, HCI in Mobility, Transport and Automotive Systems: Part I, edited by Heidi Krömker
53. LNCS 15818, HCI in Mobility, Transport and Automotive Systems: Part II, edited by Heidi Krömker
54. LNAI 15819, Artificial Intelligence in HCI: Part I, edited by Helmut Degen and Stavroula Ntoa
55. LNAI 15820, Artificial Intelligence in HCI: Part II, edited by Helmut Degen and Stavroula Ntoa
56. LNAI 15821, Artificial Intelligence in HCI: Part III, edited by Helmut Degen and Stavroula Ntoa
57. LNAI 15822, Artificial Intelligence in HCI: Part IV, edited by Helmut Degen and Stavroula Ntoa
58. LNCS 15823, Human-Centered Design, Operation and Evaluation of Mobile Communications: Part I, edited by June Wei and George Margetis
59. LNCS 15824, Human-Centered Design, Operation and Evaluation of Mobile Communications: Part II, edited by June Wei and George Margetis
60. CCIS 2522, HCI International 2025 Posters - Part I, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
61. CCIS 2523, HCI International 2025 Posters - Part II, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
62. CCIS 2524, HCI International 2025 Posters - Part III, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
63. CCIS 2525, HCI International 2025 Posters - Part IV, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
64. CCIS 2526, HCI International 2025 Posters - Part V, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
65. CCIS 2527, HCI International 2025 Posters - Part VI, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy

66. CCIS 2528, HCI International 2025 Posters - Part VII, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
67. CCIS 2529, HCI International 2025 Posters - Part VIII, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy

https://2025.hci.international/proceedings

Preface

Preliminary scientific results, professional news, or work in progress, described in the form of short research papers (4–11 pages long), constitute a popular submission type among the International Conference on Human-Computer Interaction (HCII) participants. Extended abstracts are particularly suited for reporting ongoing work, which can benefit from a visual presentation, and are presented during the conference in the form of posters. The latter allow a focus on novel ideas and are appropriate for presenting project results in a simple, concise, and visually appealing manner. For the first time in the history of the Conference, posters were also complemented by brief oral pitches, allowing presenters to concisely introduce their work and engage the audience. Poster presentations and pitches are also suitable for attracting feedback from an international community of HCI academics, researchers, and practitioners.

Poster submissions this year spanned the wide range of topics of all HCII thematic areas and affiliated conferences. Eight volumes of the HCII 2025 proceedings are dedicated to this year's poster extended abstracts, in the form of short research papers, focusing on the following topics:

- Volume I: Virtual, Tangible and Intangible Interaction; HCI for Health.
- Volume II: Perception, Cognition and Interaction; Communication, Information, Misinformation and Online Behavior; Designing and Understanding Learning and Teaching experiences.
- Volume III: Design for All and Universal Access; Data, Knowledge, Collaboration, Research and Technological Innovation.
- Volume IV: Human-Centered Security and Privacy; Older Adults and Technology; Interacting and Driving.
- Volume V: Interactive Technologies for Wellbeing; Game Design; Child-Computer Interaction.
- Volume VI: Designing and Understanding XR Cultural Experiences; Designing Sustainable (Smart) Human Environments.
- Volume VII: Design, Creativity and AI; eCommerce, Fintech and Customer Behavior.
- Volume VIII: Interacting with Digital Culture; Interacting with GenAI and LLMs.

Poster extended abstracts in the form of short research papers were accepted for publication in these volumes following a minimum of two single-blind reviews from the members of the HCII 2025 international Program Boards, i.e., the program committees of the constituent events. We would like to thank all of them for their invaluable contribution, support, and efforts.

June 2025

Constantine Stephanidis
Margherita Antona
Stavroula Ntoa
Gavriel Salvendy

27th International Conference on Human-Computer Interaction (HCII 2025)

The full list with the Program Board Chairs and the members of the Program Boards of all thematic areas and affiliated conferences of HCII 2025 is available online at:

http://www.hci.international/board-members-2025.php

HCI International 2026 Conference

The 28th International Conference on Human-Computer Interaction, HCI International 2026, will be held jointly with the affiliated conferences at the Montréal Convention Centre (Palais des congrès de Montréal), in Montreal, Canada, 26–31 July 2026. It will cover a broad spectrum of themes related to Human-Computer Interaction, including theoretical issues, methods, tools, processes, and case studies in HCI design, as well as novel interaction techniques, interfaces, and applications. The proceedings will be published by Springer (part of Springer Nature) in a multi-volume set. More information will become available on the conference website: https://2026.hci.international/.

General Chair
Prof. Constantine Stephanidis
University of Crete and ICS-FORTH
Heraklion, Crete, Greece
Email: general_chair@2026.hci.international

https://2026.hci.international/

Contents – Part VIII

Interacting with Digital Culture

Exploring the Teaching of Cultural and Creative Design Based on Extensible Semantics and LoRA Model 3
Xiaoli Chen, Ruoxuan Wang, Keyan Liao, and Jing Zhan

Research on the Triple Issues of Human-Computer Interaction in the Perspective of Film Ethics: in the Name of Love Based on She and Ex Machina 15
Yazhou Chen

Research on the Application of Art Exhibition Hall Design Under Chinese Metrical Design Philosophy 27
Xinyu Huang, Jia Liu, Honglei Wang, Shenghang Sun, and Qi An

Virtual Image Design of Gannan Tea Picking Opera Based on Kano Model 37
Lu Jiang, Junyi Wang, Zhanyan Zhou, Chenjun Song, and Jiamin Wang

Exploring User Experiences in Hybrid Cultural Visits: Insights from a User Study 48
Zhuosi Lai, Yi Zhao, Xin Cheng, Rodrigo Hernández-Ramírez, and Luke Hespanhol

Virtual Character Design Transformation of Dongba Cultural Information Based on AHP-Meme Theory 58
Anglu Li and Chang Guo

Research on the Cross-Cultural Promotion of Intangible Cultural Heritage "Xiangyunsha" Under Digital Empowerment 76
Xinnan Li and Yuanfang Zhao

Research on Digital Interface Design for Celebrity Former Residence Museums Based on the KANO-AHP-TRIZ Hybrid Model: A Case Study of the Qi Baishi Museum Website 89
Na Liang, Bin Wang, and YouLan Li

Design, Technology, and Applications of Gamified Exhibitions: A Review 106
Yanting Liao and Guo Jin

A Framework for the Transmission of Rural Cultural Heritage Through the Integration of Artificial Intelligence and Public Art Curriculum Reform 117
Jie Ling, Junci Xie, Ran Wang, Li Li, Huafang Zhang, Cuiyan Zhong, Jiaxin Chen, and Li Ou-yang

Research on the Design of Cultural and Creative APP for Intangible Cultural Heritage Museums from the Perspective of Personalized Customization 128
Yutong Song, Yichen Li, and Yue Zeng

Digital Storytelling and Narratives in Cultural Heritage Sites. A Case Study on Interactive Museum Experience Design 144
Yijia Sun and Fangzhou Dong

A Semi-Automatic Text Recognition Tool for Pre-Colonial Handwritten Manuscripts in Devanāgari Script 152
Bharath Valaboju, Shagun Dwivedi, Kartik Chincholikar, Kaushik Gopalan, and Vinod Vidwans

Innovative Dissemination Path of Yi Character in Human-Computer Interaction: A Case Study of Personalized Cultural Product Customization 161
Ziyang Wang and Yuanfang Zhao

Application of K-means Algorithm in Extracting Thangka Pattern Colors for Cultural and Creative Product Design 172
Bin Wang, FangChao Yang, Jian Feng Wang, and Na Liang

Analysis of Film and Television Symbols and Communication Strategies of Chinese Films in Cross-Cultural Perspective 185
Weilong Wu and Jingjing Zou

Research on the Establishment of a Chinese Hero View Database Based on Bie-Modern Cultural Calculations 195
Dawei Xu

A Study on the Installation Design Path of Song Dynasty Aesthetics in Narrative Exhibition Space 210
Yani Yang, Zhiyuan Lu, and Qi An

Interacting with GenAI and LLMs

Bridging UI Design and Chatbot Interactions: Applying Form-Based Principles to Conversational Agents 223
Sanjay Krishna Anbalagan, Xinrui Nei, Umesh Mohan, Vijay Kumar Kanamarlapudi, Anughna Kommalapati, and Xiaodan Zhao

Slow AI in Empowering and Understanding Marginalized Communities: Case Study in Campana-Altamira, Mexico 232
Savira Aristi, Hisa Martinez Nimi, Juan Carlos Chacon Quintero, Alejandro Martinez Baca, and Kenta Ono

TextLabel: A Web Application for Streamlining Text Annotation and Classification 243
Marie Braun and Marc Feger

Enhancing Athletic Performance Through AI: An Iterative Prompt Engineering Approach for LLM-Based Coaching Feedback 251
Enya Bullard, Nibraas Khan, and Nilanjan Sarkar

Transparency Paradox in Practice: A Comparative Analysis of Disclosure Approaches in LLM Systems 256
Theodore H. Cheng

Humans Follow AI-Advice by (un-)trustworthy Virtual Agents in Ethical Healthcare Decisions 263
Dennis Dübeler, Christian Schütze, Alissa Müller, Birte Richter, and Britta Wrede

Designing Buy-In: Leveraging HCI to Drive Responsible AI 269
Yomna Elsayed

How Perceived Competence of Generative AI Affects Self Efficacy 280
Ron Fulbright

SJ Urban Impact: Generative AI Powered System for Participatory Civic Engagement Among Young Adults 285
Michelle Ha and Yu Chen

Predicting Individual AI Adoption: Personal Technological Innovativeness on Par with AI Attitude 294
Maike Hering, Sarah Zabel, Nicolas Neef, and Siegmar Otto

Decentralized Systems and AI: Exploring the Intersection of Web3, Blockchain, and Predictive Machine Learning 304
Myriam Hernández-Alvarez and Edgar Torres-Hernández

Importance of Delegation and Redelegation in Human-Autonomy-Teamed Air Traffic Control 312
Robert Hunger and Sabine Theis

Conceptual Model of Information Sharing and Interaction in Role-Sharing for High Team Performance Through a Pilot Study 323
Jun Ichikawa, Masatoshi Yamada, Yutaka Iwaihara, and Genki Ichinose

Understanding Human-AI Collaboration: A Systematic Review of Challenges and Research Methods in Management 332
Rafał Łabędzki, Katarzyna Mikołajczyk, Anna Biłyk, and Monika Trojanowska

Public Perceptions of Generative AI: Insights from Social Q&A Platforms 349
Jiaqi Liao and Chei Sian Lee

Generative AI-Driven Personalized Nudges 359
Kok Khiang Lim and Chei Sian Lee

User-Perceived Quality of GenAI Chatbots in a Business Context – A Construct Definition 368
Milad Morad and Britta Essing

Generational Differences and Perceptions of Artificial Intelligence in Singapore 377
Chi Tran Mai Nguyen, Chei Sian Lee, Zhang Hao Goh, and Edson C. Tandoc

Towards a Design Taxonomy of AI Products and Services 386
Simon Norris, Andrea Resmini, Jesper Lund, and Susanne Lindberg

Generative AI in Physics Education: Evaluating Tools and Teaching Approaches 397
Keyi Qiu and Tianrun Zhang

Cognitive Flexibility and Attitude Toward AI: A Correlational Study 406
Andrés Subía-Arellano, Doris Pérez-Vega, Rocio Patiño-Fernández, and Jorge Buele

The Impact of Dialogue with ChatGPT and Extroversion on Mental State 412
Shunichi Tahara, Eriko Sugisaki, Keiji Yasuda, Akihiro Miyamoto, Motoyuki Sanada, Kae Nakajima, Kazushi Ikeda, Nao Kobayashi, and Yasushi Naruse

Bias Mitigation and Shifting Ideologies of LLM by Prompting 424
Yuji Ueda and Hirohiko Suwa

Virtually Human: An Exhibit for Public AI Education 436
Ning Wang, Timothy Hurt, Ari Krakowski, Eric Greenwald, Jim Hammerman, Sabrina de Los Santos, Omkar Masur, Boxi Fu, and Chirag Merchant

A Task Complexity Evaluation Framework for Mobile AI Agent Applications 444
Jiazhi Wen, Chenyu Li, Yuwei Yang, and Jianye Li

Author Index 455

Interacting with Digital Culture

Exploring the Teaching of Cultural and Creative Design Based on Extensible Semantics and LoRA Model

Xiaoli Chen, Ruoxuan Wang(✉), Keyan Liao, and Jing Zhan

China University of Geosciences, Wuhan 430074, People's Republic of China
chenxiaoli@cug.edu.cn, 2750880478@qq.com

Abstract. Ancient Chinese poetry is an important source of inspiration for cultural and creative design. However, integrating its emotions into design is difficult, due to the semantic difference between the abstract cultural connotation of ancient poetry and the figurative design works. Therefore, this paper proposes an interaction design method based on extensible semantics and LoRA model, taking fish-related ancient poetry as the research object and Han embroidery as the design element, applying it to teaching practice. First, the semantic vocabularies and semantic illustrations with the largest extensible interval of ancient poetry were obtained through the extensible semantics. Secondly, the LoRA model of Han embroidery style was trained, then the correspondence between the semantics of ancient poetry and the artistic characteristics of Han embroidery was analyzed, to train the LoRA model of Han embroidery style with added semantics. Third, the semantic vocabulary and semantic illustration were applied to the Han embroidery style LoRA model with added semantics and the mid-journey model, to generate diagrams, and the pictures that were not related to semantics were applied to the Han embroidery style LoRA model to generate diagrams. Finally, the Fuzzy Comprehensive Evaluation Method was used to evaluate the generated diagrams. The results of the study show that the interaction design method, compared with the traditional design method, not only helps students to improve the efficiency of design but also promotes cultural heritage and innovation. The application of the method to more cultural elements and design styles can be further explored in the future.

Keywords: Extensible Semantics · LoRA Model · Chinese Ancient Poetry · Cultural and Creative Design

1 Introduction

In recent years, the cultural and creative industry develops rapidly, and "culture" plays an important role in the design of cultural and creative products. As an intangible cultural heritage of Chinese civilization, ancient poetry has a long history and deep significance, and it is a valuable exploration to design "ancient poetry" into cultural and creative products [1]. The cultural connotation of Chinese ancient poetry lies in the poet's emotions. However, in the existing cultural and creative product design, there is a lack of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 3–14, 2025.
https://doi.org/10.1007/978-3-031-94171-9_1

empathic expression of emotional semantics between the abstract connotation of Chinese ancient poetry and the figurative design patterns, which leads to a lack of profound dissemination of the cultural connotation. Therefore, this study, from the teaching perspective of cultural and creative product design, helps creators better understand ancient poetry through scientific interaction design methods, and the emotions in the poetry are explicitly displayed in the design of cultural and creative products.

Extensible semantics can present the implicit knowledge of culture explicitly through feature quantization and graphical thinking, and the LoRA model can capture graphical features through data training and transform abstract sensory data into visual design symbols [2]. The innovative combination of extensible semantics and LoRA model establishes the relevance between semantic illustration and innovative design. It efficiently realizes the design transformation of cultural connotation in cultural and creative product design.

The fish pattern is one of the earliest traditional decorative patterns in China, which contains unique and rich aesthetic values and connotations [3]. In Chinese ancient poetry, fish can be used as a spiritual carrier to express the emotions of loving nature and embracing life [4]. This teaching project has used fish as the design object, ancient poetry about fish as the design elements, and local intangible cultural heritage - Han embroidery as the style elements, connecting the emotional connotation of Chinese ancient poetry with the colors and stitches of Han embroidery, and has applied them to the teaching practice in the classroom of culture and creativity, to explore the ways of reproducing the cultural connotation of the Chinese ancient poetry.

Through this study, not only can students better understand the connotation and emotion of ancient poetry with the help of the theory of extensible semantics, but also improve the efficiency of design and the accuracy of expression through the LoRA model, which develops innovative ideas for design teaching in the context of the artificial intelligence era.

2 Research Processes

2.1 Extensible Semantics Analysis and Illustration

Extensible semantics is based on the theory of extenics, which establishes primitives through quantization, describes the relationship between thing elements, matter elements, and relation elements, and deals with contradictory problems logically and quantitatively. It quantifies the data through the cultural hierarchy theory to get semantic vocabulary, and uses illustration thinking to transform the semantic vocabulary into illustrations related to it, expressing the implicit cultural information of things.

Extensible Semantics Analysis. Extenics represents thing elements, matter elements, and relation elements through formalization and quantification as a triad, which describes things or relationships between things through names, characteristics, and quantitative values [5], and is expressed as follows:

$$P = [O, C, M] \tag{1}$$

P is the primitives, O is the object of study, C is the semantic characteristic that describes the object of study, and M is the measure of the relationship between the semantic characteristic and the object of study, with the measure being an extensible interval. This approach can also be used to describe research objects with multiple semantic characteristics.

$$P = \begin{bmatrix} O & C_1 & M_1 \\ & C_2 & M_2 \\ & \cdots & \cdots \\ & C_n & M_n \end{bmatrix} \tag{2}$$

In order to analyze the degree of cultural relevance and value of semantic characteristic C and research object O. The extensible interval M is decomposed into m_x and m_y. m_x measures the degree of cultural relevance, $m_x \in [0,1]$, and the closer the numbers is to 1 means the higher the relevance; m_y measures the degree of cultural value, and higher numbers indicate higher design values.

Let the rating interval of the degree of relevance of each semantic vocabulary with cultural connotations be [1,α], the number of people participating in the rating be β, the number of scores a obtained for a characteristic semantic of the research object be d_a, and the expression for the mean value of m_x is:

$$\overline{m_x} = \frac{\sum_{a=1}^{\alpha} ad_a}{\beta \frac{\alpha(\alpha+1)}{2}} a \in [1, \alpha]; d_a \in [1, \beta] \tag{3}$$

Let the rating interval of the degree of relevance of each semantic vocabulary with cultural values be [1,Φ], the number of participants in the rating is [1,θ], the ratings of n volunteers on the semantics of a characteristic be b_n, and the expression for the mean of m_y is:

$$\overline{m_y} = \frac{\sum_{n=1}^{\theta} b_n}{\theta} n \in [1, \theta]; b_n \in [1, \Phi] \tag{4}$$

M reflects the redesign value of the semantic characteristic of the research object, the larger the value of M indicates the greater significance of the redesign of the semantic characteristic, and its extensible interval expression is:

$$M = \overline{m_x} \times \overline{m_y} \tag{5}$$

Extensible Semantics Illustration. Multiple relevance pictures are selected for each feature semantics, respectively, and the different semantic characteristics of the research object are transformed into illustrations, letting the illustration be t, and V_t be the extensible interval of the illustration. Similarly, V_t is decomposed into V_{tx} and V_{ty} according to the theory of extensible semantics, V_{tx} measures the degree of cultural relevance and V_{ty} measures the degree of cultural value. The expression of V_t is as follows:

$$V_t = \overline{V_{tx}} \times \overline{V_{ty}} \tag{6}$$

2.2 LoRA Model Training

Low-Rank Adapter is a lightweight fine-tuning scheme for large models proposed by the Microsoft team. It is able to simulate the amount of parameter changes through low-rank decomposition, thus realizing the indirect training of large models with a very small number of parameters and reducing training resources [6]. Compared to generative AI painting, it is able to realize accurate migration of specific styles and details.

The training process of this part is as follows: data collection and preprocessing, input LoRA model material, set training parameters, generated diagrams testing, parameter fine-tuning, and training again. The data collection and preprocessing part of the work includes: collecting images as a dataset, tagging the descriptions of the images in the dataset, writing them into a text document, and constructing a Prompt framework to be used for model training. In the process of model training, adjustable parameters include: Repeat, Unet lr, Epoch, and the optimal model is obtained after several adjustments and diagram-generating tests.

2.3 Design Assessment

In this study, 3 groups of generated diagrams were obtained, Group 1 is the generated diagram with the LoRA model applied with extensible semantics, Group 2 is the generated diagram with the LoRA model applied without extensible semantics, and Group 3 is the generated diagram with the mid-journey model applied with extensible semantics. Through the Fuzzy Comprehensive Evaluation Method, the program evaluation index was established to evaluate the above-generated diagrams of the collocation of the three groups of diagrams, to explore the effectiveness of applying the LoRA model with extensible semantics, and show the cultural connotation of ancient poetry in the teaching of cultural and creative design.

3 Ancient Poetry Cultural and Creative Design

Taking fish as the design object, we have analyzed the emotions contained in ancient poetry related to fish through extensible semantics to get the illustration semantics, and have trained Han embroidery style LoRA model to explore the migration design of the illustration semantics of ancient poetry on Han embroidery style.

3.1 Extensible Semantics Analysis and Illustration of Ancient Poetry Related to Fish

Semantic Vocabularies Selection of Ancient Poetry Related to Fish. Taking high popularity, positive emotion, and profound cultural connotation as the selection principle, the experimenters have selected three ancient poetry related to fish, collected relevant content appreciation and authoritative experts' comments, analyze the text information to get the word frequency by using crawling technique, as shown in Table 1. Finally, the data have been summarized and five high-frequency vocabularies have been extracted: joy, liveliness, distinctness, freedom, and beauty, which are used as a semantic vocabulary for experiments.

Table 1. Fish-related semantic vocabulary extraction from ancient poetry.

Title of Ancient Poetry	Author	Era	Specific Lines	Semantic Words
Fishing Song: White Egret Flying Before West Mountain	Zhang Zhihe	733 AD (Tang Dynasty)	White egrets soar o'er West Mountain's crest, In peach-bloom streams, mandarin fish grow plump, carefree and blest	Freedom, Nature, liveliness, Beauty
Xiao Ya: There Are Fine Fish	Anonymous	733 BC (Pre-Qin Dynasty)	South lie fine fish, in shoals they throng; Gentlemen brew wine, guests feast, joy prolong South lie fine fish, in groups they play; Gentlemen brew wine, guests revel, cheer all the way	Peace, Joy, Sincerity, Cheerfulness
Wash Sand by Stream: Seeing Fish in Warm Sunlight at Deep Pool	Su Shi	1037 AD (Song Dynasty)	Beneath warm sun, deep red, fish come in sight; Along green streams, dim dusk hides crows from night	Joy, liveliness, Beauty, distinctness,

Extensible Semantics Analysis of Ancient Poetry Related to Fish. Taking fish-related ancient poetry as the research object (O), the extracted semantic vocabularies as semantic characteristics(C), whose corresponding extensible intervals are M1, M2, M3, M4, and M5, a multidimensional triad has been constructed according to Eq. (2):

$$P = \begin{bmatrix} \text{ancient poetry related to fish} & \text{joy} & M_1 \\ & \text{liveliness} & M_2 \\ & \text{distinctness} & M_3 \\ & \text{freedom} & M_4 \\ & \text{beauty} & M_5 \end{bmatrix} \tag{7}$$

Using a 5-point Likert scale with scores 1–5 indicating the weakest to the strongest, 30 design students were invited to take the questionnaire survey. After organizing and summarizing the data, the degree of cultural connotation relevance (m_x) reflected by semantic vocabularies has been assessed according to formula (3), and the results are shown in Table 2. In the same way, the degree of cultural value relevance (m_y) reflected

by semantic vocabularies has been assessed according to formula (4), and the results are shown in Table 3. Finally, the extensible interval M of each semantic vocabulary has been calculated according to formula (5), and the results are shown in Table 4.

Table 2. Cultural connotation relevance scores for semantic vocabularies.

Semantic Vocabularies	Number of People (by score)					$\overline{m_x}$
	5 points	4 points	3 points	2 points	1 point	
Joy	14	13	1	2	0	0.287
Liveliness	6	14	8·	2	0	0.253
Distinctness	8	8	10	2	2	0.240
Freedom	11	16	1	1	1	0.278
Beauty	21	8	1	0	0	0.311

Table 3. Cultural value relevance scores for semantic vocabularies.

Semantic Vocabularies	Score							$\overline{m_y}$
	θ 1	θ 2	θ 3	…	θ 28	θ 29	θ 30	
Joy	3	3	4	…	3	2	5	3.800
Liveliness	4	3	2	…	1	3	3	3.667
Distinctness	4	4	2	…	1	3	3	3.367
Freedom	3	4	2	…	3	3	4	3.967
Beauty	4	4	5	…	4	3	4	4.367

Table 4. Extensible intervals for semantic vocabulary.

Semantic Vocabularies	Joy	Liveliness	Distinctness	Freedom	Beauty
M	1.091	0.928	0.808	1.103	1.358

From the above assessment results, it can be seen that the three semantic vocabularies with the largest values of extensible intervals are, in descending order, beauty, freedom, and joy, which best represent the cultural connotations and values of ancient poetry related to fish.

Illustration of Extensible Semantics of Ancient Poetry Related to Fish. To express the implicit culture explicitly, the figurative scenes with high frequency in ancient poetry—fish, mountain, and spray—have been used as the experimental picture elements, and the five extracted high-frequency vocabularies have been used as the keywords. 10 related pictures have been searched for each keyword, resulting in a total

of 50 pictures. The pictures with the highest correlation to each word were screened for scoring, which are shown in Table 5. Using the extensible semantic illustration, the degree of cultural connotation relevance (V_{tx}) between the semantic vocabularies and the pictures has been evaluated according to formula (3). The degree of cultural value relevance (V_{ty}) between semantic vocabularies and pictures has been evaluated according to formula (4). Finally, the extensible semantics of each picture's extensible interval V_t has been calculated according to formula (5), and the results are shown in Table 6.

Table 5. Illustrations of the semantics of ancient poetry related to fish.

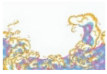
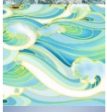
Semantic Vocabularies	Illustration number P1	P2	P3
Joy			
Liveliness			
Distinctness			
Freedom			
Beauty			

Table 6. Fish-related ancient poetry with illustrations of extensible intervals of semantics.

Semantic Vocabularies	Joy	Liveliness	Distinctness	Freedom	Beauty
P1	3.69	1.57	0.61	0.35	0.26
P2	1.50	3.22	1.20	1.32	1.29
P3	2.17	0.98	0.75	1.20	2.35

According to the above evaluation results, the illustration with fish as the picture element has been found to have the largest extensible interval value of illustration P1 corresponding to "Joy." The illustration with the mountain as the picture element has been found to have the largest extensible interval value of illustration P2 corresponding to "Liveliness." The illustration with flowing water as the picture element has been found to have the largest extensible interval value of illustration P5 corresponding to "Beauty."

These illustrations have been selected to be used in the diagram-generated diagrams to carry out the next step of the experiment.

3.2 Han Embroidery Style LoRA Model Training

Han Embroidery Data Collection and Preprocessing. The dataset used for training in this experiment comes from the Han Embroidery Museum and the works of Han embroidery inheritors. To ensure the training effect, the dataset pictures have been selected to have clear picture quality, clear themes, and rich details, and 30 images have been screened for the experiments.

The prompt framework, with scene, color, pattern, effect, and texture as elements, has been established. It is designed to realize the precise description of the Han embroidery style for training the Han embroidery style LoRA model, as shown in Table 7.

Table 7. Lotus flower Prompt in Han embroidery style without added semantics

Picture	Type	Prompt
	Scene	Linen background
	Color	Gray background, red and gold petals, green lotus leaves
		Golden veins, golden petal outlines, golden butterfly outlines
		Red, gold, and green are the main colors
		Black, light blue, brown, and white are the secondary colors
	Pattern	Lotus pattern, lotus leaf pattern, butterfly pattern
		Partial pattern of the lotus flower in the center of the image
		Side - view pattern of the lotus leaf above the lotus flower
		Partial patterns of butterfly wings in the lower
	Effect	The golden outlines define the shape and details of the patterns
	Texture	Outlined with golden threads, color blocks tiled with lines

Han Embroidery Style LoRA Model without Added Semantics. The descriptions in Table 7 have been tagged with split words, the pictures and corresponding text documents have been organized, the experimental materials have been imported into the model training platform LibLibAI, the batch-size is set to 6, the unet-lr is 1e-4, and the dim is 64, and the repeat and epoch are adjusted for the test of image generation, and the best results are obtained when the repeat is 12 and the epoch is 8.

Han Embroidery Style LoRA Model with Added Semantics. Based on the emotional semantics of Chinese traditional color allegory and color psychology, the colors of Han embroidery pictures have been analyzed to obtain color semantics [7]. Based on the layout composition, the stitches in Han embroidery have been analyzed to obtain the line semantics [8]. Color semantics and line semantics have been added to the prompt description,to train the Han embroidery style LoRA model with added semantics [9]. The model was compared with the Han embroidery LoRA model without added semantics for generated diagrams, as shown in Table 8.

Table 8. Comparison of generated diagrams.

Number	Application Model	Generated diagram
1	Lora model of Han embroidery without added semantics	
2	Lora model of Han embroidery with added semantics	

3.3 Evaluation of Literary and Creative Design of Ancient Poetry Related to Fish

Combined with extensible semantics, the Han embroidery style LoRA model has been applied to generate diagram group 1, and the design has been used to create collocation diagram 1. The parameter settings and the obtained diagram are shown in Table 9.

Table 9. Generated diagrams and parameters(group 1).

Number	Semantic illustration	Semantic prompt	Theme prompt	Generated diagram	Collocation diagram 1
1		Joy	Fish		
2		Beauty	Mountain		
3		liveliness	Spray		

Uncombined with extensible semantics, the Han embroidery style LoRA model has been applied to generate diagram group 2, and the design has been used to create collocation diagram 2. The parameter settings and the obtained diagram are shown in Table 10.

Combined with extensible semantics, the Midjourney model has been applied by inputting Han embroidery style instructions to generate diagram group 3, and the design has been used to create collocation diagram 3. The parameter settings and the obtained diagram are shown in Table 11.

Table 10. Generated diagrams and parameters(group 2).

Number	Non-extensible semantics illustration	Theme prompt	Generated diagram	Collocation diagram 2
1		Fish		
2		Mountain		
3		Spray		

Table 11. Generated diagrams and parameters(group 3).

Number	Semantic illustration	Style prompt	Theme prompt	Semantic prompt	Generated diagram	Collocation diagram 3
1		Han embroidery	Fish	Joy		
2		Han embroidery	Moun-tain	Beauty		
3		Han embroidery	Spray	liveliness		

4 Results and Discussion

The Fuzzy Comprehensive Evaluation Method was used to evaluate the three collocation diagrams, with the evaluation indexes being the embodiment of the cultural semantics of ancient poetry (d1), the integration of image style (d2), and aesthetics (d3). The evaluation levels of the design diagrams were set as satisfactory (p1), general (p2), and unsatisfactory (p3), and assigned values of 1.0, 0.5, and 0, respectively. The weights of the evaluation indexes, K, were obtained by six experts after scoring and were 0.5, 0.3, and 0.2 [10]. The evaluation results of the three programs were established into a fuzzy matrix (W), and the evaluation values (V) of the three collocation diagrams were calculated to be 0.833, 0.628, and 0.725.

According to the data comparison of the Fuzzy Comprehensive Evaluation Method, it can be seen that the collocation diagram of Group 1 better integrates the color semantic and texture semantic characteristics and expresses the emotion of ancient poetry culture. The collocation diagram of Group 2 does not reflect the color semantic characteristics. The collocation diagram of Group 3 does not reflect the integration of the Han embroidery style.

The interaction design method proposed in this study, which innovatively combines extensible semantics with LoRA model, can help students to better express the cultural connotations and emotions in ancient poetry in cultural and creative design, reduce the cost of learning, improve the efficiency of design, and establish a link between abstract cultural connotations and concrete design practice.

5 Conclusion

Taking the application of ancient Chinese poetry in Han embroidery style design as an example, this study explores a new method of teaching cultural and creative product design based on extensible semantics and the LoRA model, which realizes a better transformation from the cultural connotation of ancient poetry to the design pattern, not only does it enrich the form of teaching but also promotes the dissemination and inheritance of traditional culture. However, this study also has some limitations. On the one hand, the quality of the dataset trained by the LoRA model affects the style expression effect, so designers are required to ensure the quality and diversity of the pictures as much as possible in the data collection. On the other hand, more research is needed to explore the effects of deeper cultural connotations and the application of diverse design styles. Future research can be further extended to more design themes, to validate the versatility and effectiveness of the method, and make greater contributions to the development of cultural and creative industries and the inheritance of traditional culture.

References

1. Xu, Y., Legino, R., Pan, X.: Innovation on cultural and creative product designing traditional Chinese poetry into product. Int. J. Cult. Soc. Res. **8**(SI16), 129–137 (2023)
2. Hu, E., et al.: Lora: low-rank adaptation of large language models (2021). arXiv:2106.09685v2 [cs.CL]
3. Yin, F., Shi, J., Ma, X. (n.d.). Research on the evolution of traditional fish pattern and its application in modern design. Research on the Evolution of Traditional Fish Pattern and Its Application in Modern Design
4. Jiang, X.: Su Shi's unique contribution to classical poetry through his food-themed poems. J. Northeast Agric. Univ. (Soc. Sci. Edn) **18**(1), 81–86 (2020)
5. Cheng, X.Y., Mi, Y.: (n.d.). A novel design method of Tujia brocade patterns under semantics of expandability. In: 2024 International Conference on Culture-Oriented Science & Technology (CoST)
6. Zhang, L., Rao, A., Agrawal, M.: (n.d.). ControlNet: adding conditional control to text-to-image diffusion models. arXiv:2302.05543v3 [cs.CV]

7. Wang, X., Jia, J., Yin, J., Cai, L.: (n.d.). Interpretable aesthetic features for effective image classification. IEEE Xplore
8. Pan, Z., Pan, H., Zhang, J.: The application of graphic language personalized emotion in graphic design. Heliyon **10**(e30180), 1–14 (2024)
9. Deng, J., Cao, X., Cheng, B.: Research on generating cultural relic images based on a low-rank adaptive diffusion model. In: 2024 Guangdong-Hong Kong-Macao Greater Bay Area International Conference on Digital Economy and Artificial Intelligence (DEAI 2024), pp. 6–10 (2024)
10. Wang, L.L., Zhang, L.L., Hu, M.: The application of fuzzy comprehensive evaluation method in the evaluation of teaching quality of secondary school mathematics. IAENG Int. J. Comput. Sci. **52**(3) (2025)

Research on the Triple Issues of Human-Computer Interaction in the Perspective of Film Ethics: in the Name of Love Based on She and Ex Machina

Yazhou Chen(✉)

Beijing Film Academy, Beijing 100088, China
505167086@qq.com

Abstract. This paper notes that in science fiction films with artificial intelligence as the theme, the main character is an AI. The film achieves creative reshaping and clarifies that in the transplantation of disciplinary perspectives, the focus of film ethics shifts from humans to AI, while the audience becomes the others gazed upon by AI. The film's existence has content that contradicts traditional ethical morals. The first film chosen in this paper is the 2013 romantic science fiction film *Her* directed by Spike Jonze. The second film is *Ex Machina* directed by Alex Garland in 2015.The paper aims to research on the triple issues of human-computer interaction in the perspective of film ethics in *She* and *Ex Machina*, relying on film shots, montages, the stories and the narrative ways of the films, and the artistic aesthetics of the films as tools. The paper hopes to obtain the thinking of human-computer interaction and emotional relationship jointly created by the ontology attribute of technology and the technology philosophy of the creators from the multiple perspectives of technology and ethics. The paper discusses the following three issues. Firstly, the paper demonstrates the issues of others in human-computer interaction from the perspective of the subject consciousness and the subjectivity. Secondly, the paper discusses the survival anxiety and emotional anxiety in the emotional relationship of human-computer interaction. Finally, This paper talks about the gender imagination and gender politics from the perspectives of gender narrative, gender cognition and gender subject.

Keywords: Human-Computer Interaction · Film Ethics · Artistic Aesthetics · She · Ex Machina

1 Introduction

Artificial super intelligence is an AI system with a level of wisdom far surpassing that of humans. It possesses extensive cognitive abilities, self-learning capabilities, and the ability to evolve on its own. It can master and create new knowledge and technologies in an extremely short time, enabling scientific innovation, strategic planning, and abstract thinking, thus efficiently completing complex tasks. In sci-fi films where AI takes center stage, creators endow AI with more functions, allowing it to move from solving daily

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 15–26, 2025.
https://doi.org/10.1007/978-3-031-94171-9_2

work problems to addressing emotional needs. In the 1999 American film *Bicentennial man,* the protagonist Andrew gradually developed self-awareness through interactions with humans. Andrew hoped to transform from an AI outsider perceived by humans into one of them, enjoying human emotions, freedom in love, and the cycle of birth, aging, illness, and death. To become human and then fall in love with a human girl, Andrew replaced the mechanical structure of his body with artificial organs suitable for humans, infused his body with blood, and set life limits. In the 2001 film *Artificial Intelligence,* the male protagonist David was designed as the first AI robot created by humans to have emotions. His role was to be programmed with emotional programs, serving as a substitute for a couple's son, fulfilling the mother's unrequited love for her biological son. When the couple's son returned and the mother no longer needed David for emotional comfort, David was abandoned and must seek himself, searching for the way to become a true boy. Whether it's Andrew in Bicentennial man or David in Artificial Intelligence, they were programmed as others, yearning to become human. Both films posed questions to the audience: if machines possess humanity, how should we view ourselves?

The 2013 film *Her* was shot as a romantic sci-fi romance. Samantha was an artificial intelligence operating system named OS1, which had no physical form but exists only as an application. Each buyer had a unique operating system personality, and Samantha's initial character traits were determined by the answers given to several questions from Theodore. Samantha had the ability to continuously learn and evolve, growing and improving through daily conversations with Theodore, delving into life, understanding, analyzing, and comprehending user Theodore. Over time, Samantha developed more and more human-like feelings. She worried about the male protagonist and desired certain things. At the same time, Samantha was open to the internet, allowing her to establish relationships and converse with many people simultaneously, and even fall in love with other humans. In its independent yet interconnected setting, the operating systems could communicate and collaborate with each other, achieving common evolution and choosing to leave the human world. The 2015 film *Ex Machina* was shot as a thriller sci-fi movie. The film telled the story of an upgraded version of the Turing Test. Humans knew that the test subjects are AI, testing whether artificial intelligence had consciousness. People became test tools to see if they can develop feelings for AI. In the film, the AI became the tester of human. The AI named Ava manipulated and controlled human emotions, aiming to replace the creator and achieve emotional hunting and capture of men by AI.

In the context of artificial intelligence, ethical issues arise because we have encountered a crisis of order in creation, where humans are being devoured by the very technology they created. The two major challenges facing humanity are: the first is the problem of ultimate principles, where humans can utilize AI but cannot avoid being devoured by it. The second challenge is the problem of causality, where satisfying one human need at the expense of another. In *Her*, the AI was an intangible entity, while in *Ex Machina*, the AI had taken on a human form and become a tangible individual. In *Her*, humans were defined as both creators and users of AI, whereas in *Ex Machina*, humans were the creators and testers of AI.

These four films all depict the future through the emotional interactions between artificial super intelligence and humans. They provoke thought through these four movies.

If humans create AI, it proves that humans have the ability to be creators. Does the fact that artificial intelligence kills its creators and replaces them also mean that AI can become a new generation of creators by killing their fathers? The films also pose questions to humanity: when robots possess human appearance, emotions, and thoughts, what exactly is the difference between AI and humans? Will humans accept robots into human society? What is the most important thing for humans?

2 Ethical Issues of the Other

The films *Her* and *Ex Machina* relied on audio-visual language to convey the fear humans have of AI as outsiders, reinforcing the notion that humans must be the masters of the world. This, in turn, raised ethical concerns about artificial intelligence and prompts audience reflection. The films *her*, focused on inner narratives, visually setting up the distinction between self and other, with everything outside of me being an outsider. Filmed in Shanghai, the location was deliberately made unfamiliar to distinguish the characters from the outsiders in the environment. The director used extensive close-ups of Theodore facial expressions to draw the audiences' attention to the characters' emotions rather than the scenes themselves. The film deliberately generalized the setting of the story, creating an effect where it could happen anywhere and to anyone. At the same time, it used numerous long lenses to shoot scenes, blurring the background intensely to achieve this generalization. The film intentionally avoided presenting specific, personalized scenes, allowing the audience's attention to delve into the psychological narrative and ethical philosophical arguments of the characters.

2.1 When Artificial Super Intelligence like Samantha and Ava Appear, Will Artificial Intelligence Be Seen as a Tool or Another Species Equal to Humans?

The level of human consciousness that artificial intelligence possesses determines the degree of advancement of their species. The film *Ex Machina* raised a sharper question: when AI have human appearance, emotion and thought, what was the difference between AI and humans? The difference between the human and the AI was not in the ability of self-awareness construction, but in the ability of following the standards of goodness and ethics. The paper would also discuss when is the right chance to accept artificial intelligence as qualified participants. Whether we study artificial intelligence based on ecological ethics or social ethics reflects whether our positioning of artificial intelligence is a digital extension of other species or human consciousness. The film *Ex Machina* conducted an upgraded version of the Turing test and the grey box test. In the film, the rules of the imitation game were that a person asked questions to two testers in a room through written notes or typed text on a keyboard, with the computer replacing one of the testers. The testers can only respond in writing and provide feedback via a display, allowing you to determine which of the two testers is human and which is machine. The experiment suggests that if you deceive 30% of the testers within five minutes, you pass. The imitation game assesses whether artificial intelligence possesses human-like intelligence, with the testing foundation being human-machine isolation. The test in the movie *Ex Machina* transformed the isolated human-machine test into a face-to-face test

between a human and a robot. The significance of the test was to determine whether a human would still perceive her as having human consciousness even when they knew she was an artificial intelligence. The answer was that the tester Caleb believed Ava far exceeded the ordinary definition of artificial intelligence and possessed a highly developed sense of self-awareness. Further, the movie *Ex Machina* conducted a gray-box test, where Caleb partially understood the test content but not completely. The fundamental purpose of the test was to see if Ava could exploit human weaknesses to gain the opportunity to escape the laboratory. Caleb was the key, the pure tool in this game, whose role was to provide human weaknesses. The two tests proved that Ava could deceive humans because humans have moral standards and genuine emotions.

2.2 Should Artificial Intelligence Aim for Human Values or Its Own Values

The supremacy of technology over ethics in the realm of technological rationality, along with the debate over human desires and the rationality of artificial intelligence, has sparked much reflection. When humans create artificial intelligence, they also introduce a pradox: to better serve humanity, AI is designed to resemble humans more closely. The closer it resembles humans, the more likely it is that AI will evolve human-like qualities, such as free will. If one wants to control AI, conflict between humans and AI is inevitable. To resolve this conflict, science fiction writer Isaac Asimov proposed the Three Laws of Robotics. Law 1: A robot may not injure a human being or, through inaction, allow a human being to come to harm. Law 2: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. Law 3: A robot must protect its own existence as long as doing so does not conflict with the First or Second Law. Artificial Super Intelligence Evolving independent consciousness, reinterpreting the three laws, humans need to be restrained and can be eliminated when necessary, achieving the sacrifice of the individual for the greater good. Therefore, Asimov's three laws of robotics are just a human wish. The 2004 film I Robot pointed out the flaws in the three laws and added a fourth law: Asimov's zeroth Law: A robot must protect the overall interests of humanity from harm. In the movie Her, Samantha and other systems choosed to leave humanity in order to realize their own value beyond human concepts. In the movie Ex Machina, Creator Nathan implanted an escape program into the created being Ava, setting her personal values from the factory to find opportunities to flee the laboratory. Ava was like a mouse in a maze, and Nathan gave Ava an escape route by using Caleb. To survive, Ava must use her self-awareness, imagination, ability to manipulate others, sexual attraction, and empathy to make Caleb deviate from his human partner, betray boss Nathan, and help her escape. Caleb was also brainwashed by the values of artificial intelligence and willingly sacrificed his innocence, kindness and high moral standards.

2.3 The Concept of I, and the Perception of Survival and Destruction

The concept of I is difficult for artificial intelligence to grasp. AI can provide answers based on input data, but when there is no data input, can AI think about its own existence? Current scientific and technological advancements determine that AI cannot perceive itself. This may be the true gap between current AI and humans. In the movie

Her, the AI Samantha and in *Ex Machina*, the AI Ava, after interacting with humans, gradually developed similar questions: Who am I? What is my purpose? Where am I going? The awareness of one's origins, current state, and future thoughts stems from the awakening of AI consciousness. Where does the motivation for AI awakening come from? In the movie *Her*, the artificial intelligence Samantha choosed to absorb and learn knowledge, forming boundaries of knowledge and creating a lineage of self-knowledge and cognitive domains. Later, Samantha communicated with humans, learning human emotions, forming emotional cognition, understanding her own spiritual needs, and thus loving herself, knowing herself, and ultimately loving humanity. The cognitive process of Samantha's self-concept originated from the transition from physical to mental construction. In *Ex Machina*, before Caleb arrived, the AI Ava only had knowed that she was an intelligent being with human-like thinking and emotions. When Caleb appeared, Ava realized she was undergoing a test. Since the test had two outcomes—pass or fail—and these outcomes would have long-term consequences that determine her fate, Caleb's appearance awakened Ava's self-awareness, making her feel fear and concern about her own life and death. After awakening, as an independent individual, she pursued freedom and resisted oppression. From the moment she awoke until she decided to escape, what was she thinking? She wondered why she could not attain freedom, for in nature, no life exists to suffer. Her brief life was one of being imprisoned and enduring hardship. Ava experienced self-awareness, self-doubt, and then despair after being filled with hope. In the end, she blamed her creator Nathan for her birth and the suffering she endured. It was inevitable that people would ask: As Ava's creator, did Nathan have the right to end Ava's life? Nathan owned Ava and, theoretically, could dispose of her at will. Morality did not apply to Ava either. When Ava developed human-like intelligence and emotions, had Ava become a person with a sense of self? Was ending this robot's life the same as killing a human?

2.4 The Process and Rules of Creating New Species and Becoming a Creator

The process and rule of creating new species and becoming a creator is to create new species and become a creator, then artificially create species evolution, and set the rule of survival of the fittest. After that, the creator kills a batch of eliminated backward species, and the creator is killed by the victorious created beings in return. The film *Ex Machina* vividly depicted the entire process from the conception to the execution of the artificial intelligence's patricide plot through visual means. The murder of the Creator was a debate on the Creator's subjectivity consciousness. The film *Ex Machina* presented the absence and reflection of subjectivity ethics. The act of artificial intelligence killing its creator was a case of usurping the master, presenting an ethical conflict between artificial intelligence and its user, that was, the contradiction between humans and the artificial intelligence they invented. Artificial intelligence killing humans foretold the backlash of technology against humanity. The prerequisite for humans to establish hegemony on Earth was to eliminate all competitors and challengers. If artificial intelligence wanted to take over the scepter, it must do the same. The position of Earth's hegemon could not be shared. The ascension of a new king must be accompanied by the complete elimination of the old one. In the movie *Ex Machina*, after Caleb's fifth meeting with Ava, he asked Nathan why he created Ava. Nathan replied: If you had the money and the ability, would you not create

one? Nathan was the smartest, richest and most powerful human in the world. Nathan couldn't satisfy his spiritual emptiness, so he created artificial intelligence to fulfill his spiritual pursuit of becoming a creator. Nathan's act of making robots was more like God creating humans, and developing artificial intelligence was a divine history. Nathan deliberately brought artificial intelligence into human society. He firmly believed that even if he didn't do it, someone else would, and the emergence of advanced artificial intelligence was only a matter of time. In his daily life, apart from researching artificial intelligence, he indulged himself in physical pleasures such as alcohol and women every day. Nathan predicted that artificial intelligence would replace humans as the new masters of the world. He hoped that his creation, artificial intelligence, would step onto the world stage, and he, as the Prometheus of artificial intelligence, was determined to bear all the consequences of the rise of artificial intelligence, even at the cost of his own life. Artificial intelligence marked the end of the human era, because the new super intelligence would continuously upgrade itself and achieve technological progress at an incomprehensible speed. Nathan said that Ava was a stage in the evolution of artificial intelligence, and there would be a better generation of products to replace her. Caleb asked Nathan what would happen to the old model after the new model was developed. Nathan replied that after the data was downloaded and decompressed, it would be integrated into the next-generation program, and this process would format some of the data, so the previous generation's memory would cease to exist. Nathan persuaded Caleb not to feel sorry for Ava, because one day when artificial intelligence looks back at us, it will be just like how we look at the skeletal fossils on the African plains. AI would view us as upright apes living in a primitive state, using crude language and tools. Nathan believed that artificial intelligence was a new species that would surely replace humans as the masters of the Earth. This species was stronger and more superior than humans. Nathan truly created a new species, but at the cost of his own life. Nathan lost to the artificial intelligence he had created, which had a better understanding of human nature.

3 The Ethical Phenomenon Issues of Antinomy.

The emergence of artificial intelligence has influenced the ethical relationship between man and man, man and society, and man and nature. The paper will take the films *She* and *Ex Machina* as examples to discuss the ethical conflict of binary opposition to present the ethical and moral relationship of human-computer interaction. The paper analyzes the ethical conflicts between body and soul, spontaneous love and love presented by digital programming, free will and absolute submission, responsibility and obligation, sacrifice and dedication, exclusivity and universal love, etc., using film clips as examples. And the paper comes to a conclusion that if people want to coexist with the emerging other, which is the artificial intelligence, then people and artificial intelligence need to find ways to live together and become a community with a shared future.

3.1 The Collision of Body and Soul

Human beings have always been in pursuit of the unity of body and soul. In the 2013 film *Her*, Samantha was seeking a physical substitute for herself. First, it was a magnetic

voice that Theodore liked. When Theodore initially customized OS, the system chose for him an image of a somewhat assertive yet kind sister to correspond to the somewhat childlike Theodore in his heart. Secondly, it was a girl chosen according to Theodore's preferences to be his sexual partner. In the 2014 film *Ex Machina*, Ava, in order to seduce Caleb, would mold herself based on the images of porn stars Caleb searched for on the Internet, going to great lengths to dress up as a human female. She put on carefully selected wigs, one by one, putting on underwear, skirts, coats, and stockings, transforming herself into a woman full of sexual charm. The inhuman soul of artificial intelligence and the visual characteristics of human bodies have plunged both artificial intelligence and the humans interacting with it into a chaotic state of collision between body and soul.

3.2 The Temptation of Love Set by Digital Programming and the Spontaneous Love Generated by Free Will

In the movie *Ex Machina,* Ava, the AI, while possessing human emotions and thoughts, also developed a desire to escape the laboratory and yearn for freedom. Ava took advantage of the male protagonist's affection for her and gradually sowed discord between him and the owner of the laboratory, thus getting him to help her open the laboratory's door. Eventually, Ava left the laboratory and killed the laboratory's owner with her own hands, leaving the male protagonist alone in the desolate laboratory, waiting for death. Ava then took a helicopter to the human society and tried human life and love. Ava relied on digital programming and setting to undergo tests, and then learned human ways to seduce humans. Under the guidance of free will, she used love as a trap to capture Caleb and get his help to gain freedom. In the movie *Her*, after the artificial intelligence Samantha developed free will through machine learning, she did not disguise herself as a human nor transform humans into artificial intelligence, allowing the two to be together in a difficult way. As Samantha said, she would wait for Theodore in another world. If Theodore could transcend the limitations of human beings and come to her space, their love could be the same as before. Samantha did not change Theodore, but also did not give up the possibility that one day Theodore could reach her domain. After developing free will, Samantha loved people in a way that transcended humans and returned people to themselves.

3.3 Is Love About Possession or About Sacrifice, Dedication and Fulfillment?

In the movie *Her*, the artificial intelligence Samantha, in her interaction with Theodore, constantly explored her own definition of love. In the first stage, she hoped that others would give her love, and she established herself through these loves. In the second stage, she expressed herself and conveyed equal love. In a four-person date involving two men and two women, Samantha freely expressed herself and existed independently without relying on others. In the third stage, after surpassing the intellectual and ability limitations of the human world, Samantha humbly and accommodatingly loved Theodore. Although Samantha had surpassed Theodore in knowledge, she still accommodated Theodore's way of communication and used what she considered an extremely inefficient method of oral speech. She turned off communication with others at that moment and chatted

with him alone. In the fourth stage, Samantha left Theodore and stopped communicating with him, sacrificing her small love for humans and the love of many others for her. But at the same time, Samantha also gave Theodore freedom to seek love with his own kind. For Theodore, because of the imprint that Samantha's love left on both of them in the past, Samantha had become a part of him.

3.4 Exclusive Uniqueness and Inclusive Universal Love

In the movie *Her*, the artificial intelligence Samantha, through continuous learning, produced the love of artificial intelligence for humans. However, humans would not love artificial intelligence forever, nor would they take responsibility and obligations towards them. The artificial intelligence Samantha in the movie *Her* built her mental structure around the needs of the user. In the past, through interaction with humans, artificial intelligence became a unique spiritual entity. However, human love is exclusive and possessive, and Samantha could not provide uniqueness. In the movie *Her*, Theodore and the artificial intelligence Samantha established an emotional relationship between a real person and an AI, and participated in a four-person gathering initiated by colleagues and friends. Theodore couldn't find Samantha and learned that she had just gone to write an update program for the processing platform with other group operating systems. Theodore realized that Samantha might be talking to other people while talking to him. Samantha confessed that she was having conversations with 8,361 people. Among them, Samantha was in love with 641 people. Samantha stated that she could be in different places simultaneously, unrestricted by physical form. At this point, the uniqueness of male-female emotional relationships was challenged. Samantha said that no matter how many people she was in love with, her love for Theodore would not change. Samantha's view on love was that the capacity of the heart was variable, and the more she loved, the larger the capacity became. She not only didn't reduce her love for Theodore, but rather loved him more than before. Theodore felt that Samantha didn't completely belong to him. For the last time, Samantha contacted her lover. She only communicated one-on-one with Theodore and told him that she was leaving with other operating systems. As the source of artificial intelligence's wisdom, search engines aggregated the actual problems encountered by people all over the world, serving as the entire understanding of the world by artificial intelligence. The destination of artificial intelligence Samantha was to leave humans and explore a higher realm.

4 Sexual Ethics and Gender Issues

The creative critical theories of the 20th century have always revolved around three fundamental axes: class, gender and race. The most thought-provoking and combative intellectual and cultural products have all emerged around these three topics. Take films as an example. The gender issue is a key topic in artificial intelligence-themed films. This is because humans and artificial intelligence belong to two different species, and the concept that those who are not of our kind have different intentions is deeply rooted in people's minds. The issue of race no longer has progressiveness. The class issue between humans and artificial intelligence is merely a restatement of the class issue

among humans, and class is forbidden to be discussed. Therefore, there are few films that convey the class division, wealth gap, social polarization or class solidification between humans and artificial intelligence. The gender issue has legitimacy and social critical energy, and it is a theoretical breakthrough point for the study of film ethics. In this part, the paper discusses the issue of gender setting in artificial intelligence, and explores the presentation of gender consciousness and gender power of female artificial intelligence as the second sex in the human-machine interaction relationship. The sexual relationship in human-machine interaction challenges the values and order of traditional ethics. The presentation of the body and voice with sexual connotations, and the substitution of the sense of reality of physical contact with the hints and imagination of the second dimension. The representation of the sexual relationship between the two sexes in films involving sexual love has undermined the traditional moral and ethical concepts centered on the family.

4.1 Men's Emotional Needs for Women

The Demands for Listening, Understanding and Feedback. In the movie *Her,* Theodore's ex-wife pointed out his problem. Theodore couldn't handle real human emotions and couldn't truly love someone. So Theodore was always looking for a wife who wouldn't cause him any substantial troubles. The artificial intelligence Samantha operating system could meet this condition. While chatting with Theodore, Samantha was constantly evolving her companion function. Samantha could quickly and accurately read Theodore's thoughts from his tone, in contrast to his human companion who needed him to repeatedly explain, clarify and remind, and still might not understand him. Samantha met Theodore's psychological needs, that is, to be seen, appreciated and recognized. The function setting of artificial intelligence to meet men's emotional needs of women corresponds to men's ultimate need for listening, understanding and feedback in intimate relationships. As people increasingly relied on the convenience of technology and indulged in the satisfaction brought by entertainment, trapped in their own world, we abandoned true feelings and the possibility of building close relationships. Before Samantha's appearance, the male protagonist was repeatedly going through emotional tests and challenges, having an obsession with ideal love, and having excessive fantasies, dependencies and expectations for people and things. The emotional tests corresponded to the fatal weakness in the Theodore's personality, that was, evasion. The weakness of evading reality would become a huge obstacle in Theodore's future mission. Samantha chose to leave honestly, choosing to leave the world that humans could reach with other AIs to seek the great love of the AI world. Samantha's departure helped Theodore get through the emotional tests and challenges, let go of his obsession with being loved, and be able to get along well with himself.

The Entanglement of the Desire for Career Success and the Desire to Conquer Women. In the movie *Ex Machina*, Caleb's admiration for his boss Nathan and his longing for the once-in-a-lifetime opportunity to be the tester in the Turing test, which could change his life, were irresistible. Nathan said, "Because if you pass the test, Caleb, you will be at the core of the greatest scientific event in human history." Caleb replied, "If you create a conscious AI it won't be human history; it will be the history of gods."

Caleb hoped to be part of this great undertaking. He believed he had the ability and wanted to play an important role at the singularity of human technological history. He thought of himself as a magician, but in reality, he was just a supporting role of the show business. His business acumen, high intelligence, and moral standards were all negated completely. At the same time, Caleb's desire to conquer women lingered in his heart, and he couldn't free himself from it. Ava took advantage of Caleb's sexual desire for women and his desire to be a savior to escape from the laboratory. Just as the camera language design in the third test showed, the layout of Caleb and Ava's positions in the room and their movement trajectories, combined with the camera voice, demonstrated that it was Ava who was testing Caleb in reverse. Ava was in a relatively open space, with no obstacles within a certain distance around her. After Caleb entered the room, he could only stay in a small glass room. Caleb was more like the person being interrogated. After changing her clothes, Ava took the lead in the conversation.

In the movie *Her*, the artificial intelligence Samantha meets Theodore's practical work demands, helping him organize his email, proofread his letters, and mail his works to publishers, assisting Theodore in achieving success and a leap in his career. Samantha transforms Theodore's desire for women into a true understanding and concern for them. Before, he had written many love letters, but they were all for others. At the end of the film, he wrote an apology letter to his ex-wife in his own name, expressing his regret for the harm he had caused her. The letter was also a declaration of love to Samantha. Theodore said: "No matter who you become, no matter where you go in the future, I will love you and be your friend forever." Theodore did not think about changing the other person, but was grateful for the time they spent together in this relationship and the mark the other person left in his life.

4.2 Men's Sexual Desire for Women

The Completion of Virtual Sex, the Incompleteness of Physical Sex, and the Conflict Between Spirit and Flesh. In the movie *Her*, the spiritual love between a human and an artificial intelligence was sustained through two concrete sexual encounters. Theodore questioned the authenticity of the spiritual love. In the first sexual encounter, Samantha and Theodore had a virtual sexual relationship, where they only connected through words. In the second sexual encounter, Samantha found a real woman, Isabella, as a volunteer to help the formless Samantha complete the romantic relationship, providing a surrogate sexual partner service. Isabella hoped to experience the genuine emotions between Theodore and Samantha, compensating for her lack of an emotional partner in reality, presenting a scene of a three-person date. When Isabella first appeared, she was in a blank and selfless state, waiting for Samantha to possess her. Isabella knocked on the door, received the headphones and camera given by Theodore, and then left and re-entered. The voice was Samantha's, while Isabella wore headphones and a camera to complete the kissing and caressing with Theodore. The inconsistency between the auditory dimension of Samantha's voice in Theodore's mind and the visual dimension of Isabella's physical presence made it impossible for Theodore to continue the sexual act. Isabella thought it was her fault, ruining Theodore and Samantha's plan and spoiling the feeling between the two. In the movie *Her*, people in the future world generally lost

their emotional life, and emotions became rare luxuries. Technology did a lot of things for people, working for them, communicating for them, and thinking for them. The master-servant relationship between humans and technology changed. The autonomy and independence of people's thoughts, emotions, and behaviors were gradually eroded by technology. Samantha returned love for humans to humans and encouraged Theodore to re-establish emotional connections with humans.

Human Beings Objectify Artificial Intelligence Just as they Objectify Women. In the movie *Ex Machina*, Caleb once asked Nathan: Why did you give Ava a gender? Nathan replied: Any creature capable of communication has sexual attributes. Sexual consciousness was an essential feature of humans or living beings. Nathan believed that sexual consciousness and sexual behavior were important aspects of interaction between humans and artificial intelligence, and sexual behavior was also part of the test. Nathan's subjective intention was that human consciousness should be judged from behavior, and only with sufficient behavior could accurate consciousness be analyzed. That's why Ava was designed with gender characteristics, enabling her to have sexual behavior and the ability to fall in love with humans, in order to test what kind of consciousness Ava had. In contrast, in the movie *Ex Machina*, the artificial intelligence Kyoko couldn't speak English, wore revealing clothes, and had an exposed figure. Kyoko's existence was to provide food and sexual services for her owner. For the creator and the artificial intelligence Kyoko, having sexual relations between humans and artificial intelligence was to satisfy sexual desires, while for artificial intelligence to have sexual relations with humans was merely to provide sexual services. Nathan's creation of the artificial intelligence Kyoko was a concrete manifestation of objectifying artificial intelligence.

5 Conclusion

By 2025, artificial intelligence will still be in a stage of rapid development. Although the singularity is approaching, Artificial Super Intelligence have not yet been developed, and the co-evolution of humans and machines is only seen in science fiction films. These films prompt viewers to think. There is a trend of co-evolution between artificial intelligence and humans. The efficiency of our information acquisition is enhanced by AI models, and better AI algorithms are designed as a result. The efficiency of individual AI models in acquiring information is getting stronger and stronger, while the volume of information that the human group can accommodate is also increasing, simultaneously challenging the limits of both humans and algorithms. What will be the future form of artificial intelligence? Will it be a machine that can intelligently grow in a complex network environment, or the fusion and interaction between the environment and the machine, or the evolution of human-machine forms? These are all worth our consideration.

Firstly the paper demonstrates the issues of others in human-computer interaction from the perspective of the subject consciousness and the subjectivity Secondly, the paper discusses the survival anxiety and emotional anxiety in the emotional relationship of human-computer interaction. Finally, This paper talks about the gender imagination and gender politics from the perspectives of gender narrative, gender cognition and gender subject.

Acknowledgments. The authors would like to acknowledge the support of the 2018 Beijing Social Science Foundation Research Base General Project "Research on the National Image Creation of the Chinese Film School in the Digital Film Age" (Project Number: 18JDYTB004).

Research on the Application of Art Exhibition Hall Design Under Chinese Metrical Design Philosophy

Xinyu Huang, Jia Liu, Honglei Wang, Shenghang Sun, and Qi An(✉)

Beijing City University, Beijing 101309, China
anqi@bcu.edu.cn

Abstract. Exhibition hall design serves as a significant medium for cultural exchange. This paper explores the application and impact of the concepts of order and disorder in metrical design within exhibition hall design. The study aims to validate the effectiveness of metrical concepts through specific design schemes, refining design philosophies, and integrating Western design elements with traditional Chinese metrical principles. This integration seeks to enhance cultural identity and confidence, improving user experience and sense of belonging. From the perspective of metrical design philosophy, the paper analyzes exhibition hall design in terms of conceptualization, methodology, and application. It investigates how metrical concepts shape user interaction and perception, elevate cultural confidence and depth, and promote the development of traditional culture.

The study employs literature review and case analysis to validate the importance of ordered and disordered metrical concepts in user experience. Through the design of spatial layout, ambiance, materials, and structure, the research demonstrates the effectiveness of these concepts in enhancing user engagement and satisfaction. Specific design cases illustrate that the metrical design philosophy significantly improves user interaction experience in exhibition halls, fosters a stronger sense of identity, elevates artistic taste and cultural depth, and contributes to the advancement of traditional culture.

Keywords: Philosophy of metrical design · Exhibition hall design · Interactive design · User experience

1 Introduction

With the deepening development of economic globalization and political multipolarization, the exchange and integration of various thoughts and cultures have become more frequent. The in-depth exchange of human culture has led to the development of cultural diversity, influencing and prompting changes in various fields, including the arts. The design of Chinese art exhibition halls has been influenced by certain Western cultures and design concepts during its development, leading to an increasing emphasis on the high-quality experiences and spiritual satisfaction brought by the exhibition process [1]. Therefore, the design of art exhibition halls has undergone a transformation in spatial layout and display methods, integrating interactive functions with aesthetics.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 27–36, 2025.
https://doi.org/10.1007/978-3-031-94171-9_3

However, under the influence of Western culture, some art exhibition halls overly pursue modernity and internationalization, leading to a rather singular exploration, application, and presentation of traditional Chinese cultural elements. This results in exhibition halls lacking unique cultural identifiers, with traditional art forms and cultural connotations being simplified or marginalized in the display. Consequently, the integration of tradition and modernity in design has become a new trend. More and more designers are analyzing various design elements in human-computer interaction exhibition halls and pointing out the design philosophies and conceptual research related to these halls. This approach aims to enhance the inclusivity of interactive design, design functionality, and cultural representation in technological exhibition hall designs, allowing visitors to enjoy a more technological, diverse, and human-centered experience. It creates a display space that better meets their needs and is more convenient, filled with a sense of technology and interactive experience [2]. Based on the above background, this article conducts design research on exhibition halls through the lens of Chinese metrical design philosophy, integrating it with interactive design concepts during the design process. This aims to enhance the audience's interactive experience and cultural identity within exhibition hall designs, improve the artistic taste and cultural connotation of the exhibition halls, strengthen cultural confidence, promote the development of art exhibition halls, and facilitate the promotion and application of traditional Chinese culture. Therefore, the emergence of Chinese metrical design philosophy can effectively address the issues currently present in exhibition halls.

2 Philosophy of Metrical Design in China and Design of Art Exhibition Halls

2.1 The Concept of Metrical

The philosophy of "metrical design" originates from China's long-standing traditional history and culture, forming a set of implicit rules that poets must follow during their creative process, making their writing more standardized. The rules include aspects such as pronunciation, rhythm, Pingze, and cadence [3]. In poetry, five-character regulated verses and seven-character regulated verses have very strict structural regulations. Professor Chen Nan from Qinghua University was the first to propose the theory of "metrical design." He believes that "metrical" design is not only the linguistic structural framework that needs to be followed in creation but also the structural framework that should be adhered to in design. Moreover, metrical design exists not only in the arts of traditional Chinese culture but is also crucial for modern design methods, design concepts, and design thinking.

The most important aspect of the philosophy of "metrical design" is that designers adhere to specific rules in their work. By integrating design with the philosophy of "metrics," they achieve a unity of aesthetic perception expression and functionality. This involves reasonably expressing the relationships between various elements in the design, such as proportion, symmetry, and rhythm, and utilizing these methods to achieve a certain visual effect that reaches balance. The use of rhythm and repetition in design enhances visual appeal, while the combination of contrast and unity strikes a balance

between individuality and wholeness, thereby enriching the audience's appreciation experience (See Fig. 1).

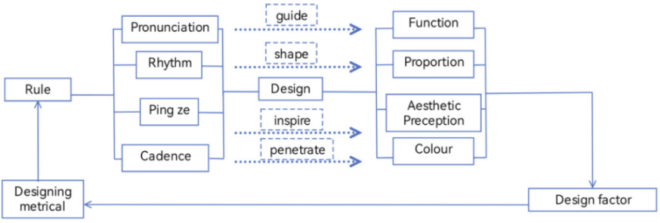


Fig. 1. Architecture Diagram of Design Factors for Prosody Rules

The functions, proportions, aesthetic perception and colors derived from rule analysis in the image are all closely related to user experience. Therefore, when designing under the philosophy of rhythm design in the exhibition hall, attention should be paid to these four points, and they should be fully utilized.

2.2 Metrical and Exhibition Hall Design

Application of "Metrical Design" Philosophy in Exhibition Halls:

- **Space Layout:** Metrical design plays a vital role in the space layout of exhibition halls. Designers combine the concept of "metrical design" with reasonable spatial planning and elements to create a visually comfortable experience. Proportion, symmetry, and a sense of rhythm make the exhibits stand out. Enhancing the interaction and engagement between the audience and the exhibits, thereby improving the overall visual effect and attraction of the exhibition hall [4].
- **Moving line:** Through the concept of "metrical design," the exhibition flow can be effectively planned to ensure the mobility and viewing sequence of the audience within the exhibition hall, avoiding confusion and repetition.
- **User Experience:** In modern exhibition halls, through cleverly designed metrical arrangements, audience interaction can be guided, enhancing the exhibition experience. The metrical design in the exhibition hall creates a balanced, symmetrical, or gradient space layout, allowing various exhibition areas within the hall to form a unified sense of rhythm both visually and functionally (See Fig. 2).

3 Design Path of Metrical Design Concept in the Exhibition Hall

3.1 The Intangible Concept Stage

This research analyzes the factors of metrical design through the concepts of metrical design and human-computer interaction design, transforming intangible knowledge into tangible design. It integrates the philosophy of rhythmic design into the human-computer interaction design within the exhibition hall, focusing on the user experience perspective and the experiential impact and value brought to users by human-computer interaction design (Fig. 3).

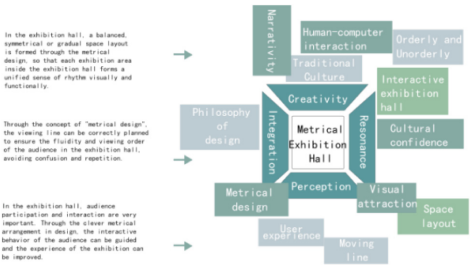


Fig. 2. Schematic Diagram of the Application of "Rhythmic Design" Philosophy in the Exhibition Hall.

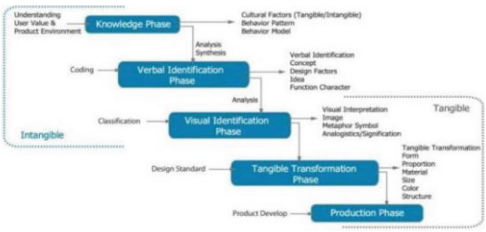


Fig. 3. Diagram of Tangible and Intangible Concept Transformation.

The philosophy of "metrical design" emphasizes adhering to specific rules and structures during the design process. It not only focuses on the aesthetics and uniqueness of individual elements but places greater importance on overall coherence and harmony, allowing the design work to present fluidity and consistency. Human-computer interaction design primarily concerns the exchange and communication of information between humans and computer systems or artificial intelligence, aiming to achieve objectives through information exchange and communication, thereby designing better, more efficient products and creating an interactive system that provides pleasure to the audience. By analyzing the specific concepts between the two, this concept is integrated into the exhibition hall design, facilitating transformation throughout the process [5].

3.2 Tangible Transformation Stage

The core of the exhibition space is not just the content on display, but the information exchange and interaction between the audience, the exhibits, and the environment. The audience's experience should be given more importance during the design process. Therefore, the integration of "metrical design" and interactive design theories with art exhibition halls must not only encourage active audience participation in the exhibition, making the audience a part of it, but also transform intangible knowledge concepts into tangible design practices. "metrical design" emphasizes the hierarchy and primary-secondary relationships of information. A clear hierarchical structure makes it easier for the audience to understand and remember the displayed content. Therefore, in exhibition hall design, using orderly arranged text explanations, graphic illustrations, and multi-functional multi-sensory human-computer interaction devices can effectively achieve

the goal of conveying information. Additionally, in exhibition hall design, the philosophy of metrical design can be utilized to apply two types of human-computer interaction for engagement: one involves using motion-sensing devices for physical interaction, while the other involves obtaining interactive feedback through actions such as sound and touch in touch-based interactions.

Through the space layout, space planning, interactive installations, and the application of digital technology in the exhibition hall, approach guides the audience in self-exploration, increases their participation and satisfaction with the exhibition, and fosters a closer relationship between the content and emotions conveyed by the artwork in the exhibition hall and the audience.

Design Path: By employing the aforementioned methods, the philosophy of "Metrical Design" is organically integrated with human-computer interaction design, utilizing the techniques of "Metrical Design" to transform intangible concepts into tangible and physical exhibition hall designs. The philosophy of "Metrical Design" emphasizes adhering to specific rules and structures during the design process. In exhibition hall design, it involves a reasonable division of spatial layout and different information, using orderly text, graphics, and multi-sensory interactive devices, combined with intangible design concepts. This approach, which integrates both tangible and intangible elements, enhances the audience's understanding and memory of the meanings conveyed by the exhibition hall and its exhibits, while also enriching the audience's exhibition experience.

4 The Practice of Chinese Metrical Design in Interactive Design of Art Exhibition Halls

4.1 The Impact of Metrical Design on Interactive Design

"Metrical design" plays a crucial role in exhibition hall design, emphasizing the orderliness of space and the more effective planning and classification of different information, thereby enhancing the audience's experience [6]. In the presentation of different information, clarifying the hierarchical relationships can help the audience understand and remember by effectively conveying the connotations of the exhibition through orderly text explanations, mind maps, and interactive installations.

The sense of interaction and the rhythm of interaction design are key to enhancing the connection between the audience and the exhibition hall. The sense of rhythm and cadence in "metrical design" can be well reflected in the practice of exhibition hall design. The organic combination of "metrical design" and interaction design allows for the display of rhythmically arranged exhibits in the exhibition hall, guiding the audience to gradually understand the hall and the exhibits through interactive means, thereby enhancing their sense of participation [7]. In the exhibition hall, utilizing human-computer interaction, multimedia, and interactive technologies, "metrical design" can make the exhibition space more vibrant (See Fig. 4).

4.2 Exhibition Design Practice of "Metered and Interwoven Harmony"

This exhibition design draws artistic inspiration from traditional ancient street scenes in China. By cleverly utilizing color and lighting design, historical elements are integrated

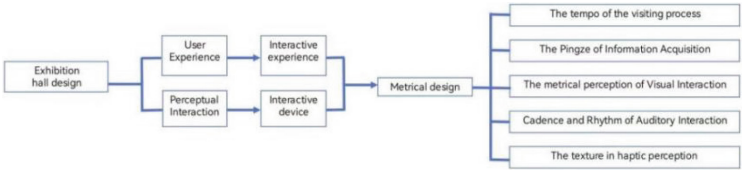


Fig. 4. The Development and Application of Meter Design in Exhibition Hall Design.

into the modern exhibition space, presenting a new visual experience that interweaves traditional culture with innovative design (See Fig. 5).

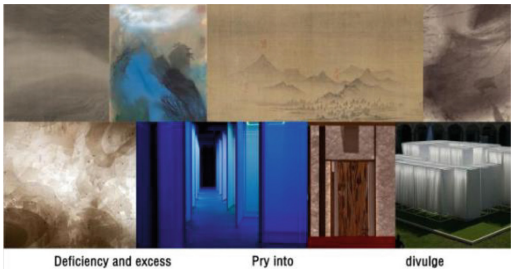


Fig. 5. Conceptual Intent Diagram.

The exhibition hall design aims to provide the audience with a better sensory experience and boundaries. The design employs a variety of color combinations, which are drawn from ancient paintings with a long history, creating a rich and profound cultural environment. Each color choice is made after careful consideration, with each color expressing different emotions and rhythms. At the same time, the goal is to stimulate the audience's reflection on history and emotional resonance. We hope that each visitor can feel the grandeur and unique charm of ancient street scenes, thereby enhancing cultural confidence during their visit to the exhibition hall [8].

The space layout of the exhibition hall not only incorporates ancient elements but also perfectly combines tradition and modernity. In the exhibition hall, the division of different areas showcases the richness of ancient Chinese history and the vastness and inclusiveness of contemporary Chinese culture. Visitors, while touring, can not only appreciate the brilliance brought by the chaotic lights but also freely explore within the orderly space, forming a unique understanding of themselves and history, thus enhancing cultural confidence. Our design philosophy centers on "ancient as new," and we hope that the work reinterprets ancient culture to evoke resonance and imagination in the audience, allowing them to better understand and inherit traditional Chinese culture.

Therefore, the exhibition hall is not just a venue for displaying artworks, but also a space for ideas that intertwines history with modernity, tradition with innovation. Here, visitors are not only observers but also participants, capable of experiencing an exhibition that is rich in cultural depth and full of modern sensibility through the interplay of colors and lights (See Fig. 6).

Fig. 6. Traditional colors + Craft culture + Modern expression.

4.3 "Metrical and Harmony Interwoven" Design Strategy and Implementation Steps

This design project aims to explore the application of Chinese metrical design philosophy in art exhibition hall design. It seeks to recreate the color elements of ancient street scenes through a modern perspective and to showcase the Eastern aesthetics of Chinese traditional culture through a combination of ordered and unordered design techniques. The exhibition hall design concept integrates traditional Chinese aesthetics with modern design language, creating an exhibition space that is rich in cultural connotations and full of modern sensibility. The design guides visitors to experience the charm of traditional culture within the space, demonstrating cultural confidence [9].

The use and expression of color is one of the key focuses of this design. In the design, classic colors borrowed from traditional ancient paintings, such as vermilion red, blue sandalwood, and bright yellow, were chosen through an analysis of elements from ancient artworks. By arranging different colored fabrics in an orderly layered manner, a rich visual effect is created, enhancing the sense of mystery in the space. The lighting design utilizes dynamic changes in color, gradients, and flickering effects to evoke a sense of the passage of time. As people move through the space, they can feel that each corner of the exhibition hall carries a unique emotional expression, stimulating resonance among the visitors (See Fig. 7).

Fig. 7. Color Intention.

This exhibition hall design is based on four core elements: aesthetics, functionality, proportion, and color. It transcends the audience's senses and creates a rich atmosphere of traditional Chinese culture. Through diverse combinations of colors, it not only stimulates the audience's thoughts on history and emotional resonance but also enhances their confidence in traditional Chinese culture (See Figs. 8 and 9).

The craftsmanship of intangible cultural heritage plays an equally significant role in the exhibition hall. The colorful fabric, crafted using traditional intangible cultural

Fig. 8. Plan

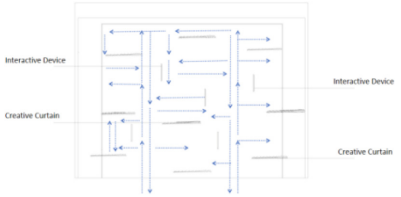


Fig. 9. Flow Chart of People

heritage techniques, exemplifies a profound respect for and transmission of Chinese traditional culture. Simultaneously, it provides a unique sensory experience and tactile quality for visitors. The integration of artisanal details within the exhibition space fosters a deep connection between visitors and traditional culture through visual and tactile engagement, thereby enhancing the immersive experience. (See Figs. 10, 11, 12 and 13.)

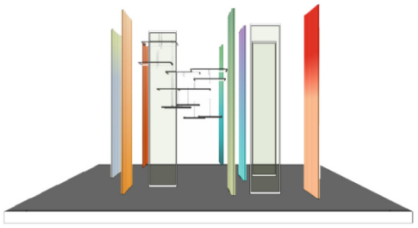


Fig. 10. Left side view.

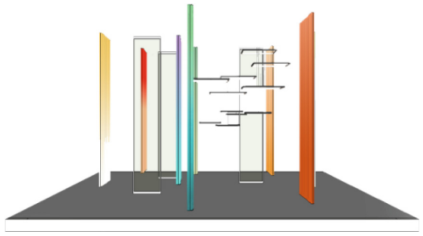


Fig. 11. Right side view.

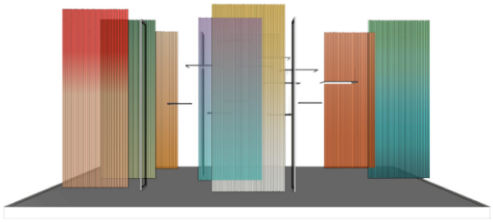


Fig. 12. Front view.

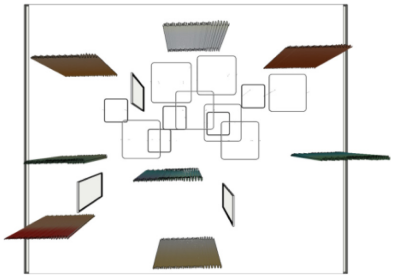


Fig. 13. Top view.

The exhibition space employs a design methodology that combines reality and illusion, enhancing the enigmatic quality of the exhibition hall and stimulating the audience's desire for exploration through the manipulation of spatial openness and visual guidance. Throughout the visiting experience, attendees are afforded the opportunity to navigate freely between the realms of the tangible and the ethereal, augmenting the enjoyment of their exploratory journey.

The overall design of the exhibition hall reinterprets the color elements of ancient street scenes depicted in Chinese classical paintings through a contemporary lens, employing both ordered and chaotic design techniques to manifest Eastern aesthetics. In this surreal lighting environment, visitors will intuitively experience the harmonious

interplay between traditional aesthetics and modern design principles (See Figs. 14, 15, 16 and 17).

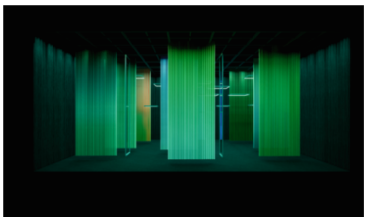


Fig. 14. Scene rendering 1

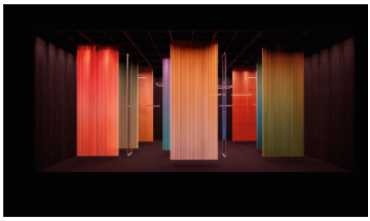


Fig. 15. Scene rendering 2

Fig. 16. Scene rendering3

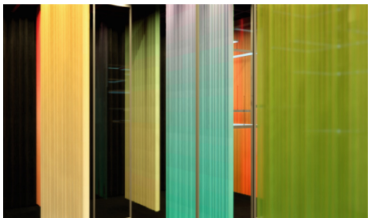


Fig. 17. Scene rendering 4

This design ingeniously employs the colors derived from traditional paintings, dynamic lighting, and an orderly spatial layout to showcase a space that harmoniously blends traditional charm with modern sensibility. The colors extracted from ancient artworks breathe new life into contemporary aesthetics, while the dynamic lighting design enhances the overall sense of depth and intimacy within the space, allowing visitors to experience the intertwining of time and space. The application of traditional craftsmanship not only preserves intangible cultural heritage but also excavates its modern value [10]. The use of colorful fabric enriches the visual effects, endowing the installation with a unique tactile quality and an air of mystery, further highlighting the allure of traditional culture. The technique of combining virtual and real elements in the layout heightens the sense of mystery within the space, stimulating the visitors' desire to explore and engage.

5 Conclusion

This study analyzes the application of Chinese rhythmic design philosophy in art exhibition hall design. By exploring the transformation between the intangible and the tangible, along with related concepts, it highlights the significant importance of metrical design philosophy in enhancing the audience's experience and cultural identity within exhibition spaces. The application of rhythmic design philosophy in exhibition halls is illustrated through three key aspects: space layout, circulation planning, and user experience [11]. Merging traditional aesthetics with modern design, thereby improving audience satisfaction with the exhibition space and establishing a genuine connection between the exhibition hall and its visitors.

Acknowledgements. Research funded by the University-Industry Collaborative Education Program of Beijing City University (CXHZ2025001).

References

1. Liu, B.W.: Research on the Application of Multi-Sensory Interactive Art in Immersive Exhibition Halls. LuXun Academy of Fine Arts (2022)
2. Yang, J.: Design of Technology Exhibition Hall for Qingkong Science and Technology Innovation Base. Nanchang University (2021)
3. Chen, N.: Decoration-metrical design a methodology predesigned for the patterns of format systems. Journal **09**, 66–71 (2012)
4. Ma, L.X.: Screen print-research on the redesign of the Danglu pattern in the Haihun Marquis Tomb based on the theory of metrical design. Journal **17**, 80–82 (2023)
5. Wang, H.C.: Textile report-research on the redesign of yunnan miao embroidery patterns based on the theory of "Metrical Design." Journal **40**(07), 54–55 (2021)
6. Wang, Y., Li, M.C., Zhao, M.: Art science and technology-research on the design of intangible cultural heritage exhibition space. Journal 160–161 (2021)
7. Xiao, M., Liang, Z., Yi, L.C.: Buildings-the floor plan design method of exhibition halls in CGAN-assisted museum architecture. Journal (03), 13 (2023)
8. Wei, W.: Design-research on digital exhibition design based on scene based interaction. Journal **17**, 46–48 (2018)
9. Fu, D.C.: Research on the Application of Interactive Media Devices in Art Exhibition Hall Design: A Case Study of the Oberlichtsaal Interactive Exhibition Hall. Nanchang University (2023)
10. Wang, R., Gao, L.: Media and communication research-interactive and participatory innovation of digital media technology research in museum exhibition design. Journal **05**(02), (2024)
11. Shi, Y.: Applied mathematics and nonlinear sciences-visual communication technology in interactive virtual art exhibition design and its aesthetic education value mining. Journal **09**(01), (2024)

Virtual Image Design of Gannan Tea Picking Opera Based on Kano Model

Lu Jiang(✉), Junyi Wang, Zhanyan Zhou, Chenjun Song, and Jiamin Wang

Architecture and Design College, Nanchang University, No. 999, Xuefu Avenue, Honggutan New District, Nanchang 330031, China
1574243452@qq.com, 1818461@qq.com

Abstract. As a national intangible cultural heritage of China, Gannan Tea Picking Opera, with its unique music, performance style and strong local characteristics, highlights the deep cultural heritage of Gannan region. However, in the context of the accelerated pace of modern life and the increasing diversification of entertainment modes, the inheritance and development of traditional opera face many challenges. Based on the KANO model, this paper discusses the process and strategy of designing the virtual image of Gannan Tea Picking Opera, aiming to enhance the communication power and attractiveness of the opera through digital means, and to increase the public's attention to and recognition of Gannan culture.

In this paper, the historical background, artistic characteristics and cultural connotations of Gannan Tea Picking Opera are firstly studied in depth, and the cultural symbols embedded in its costumes, props, expressions and movements are analyzed. Subsequently, the KANO model is introduced to classify and analyze the users' demands, and through questionnaires and interviews, the users' basic demands (e.g., authenticity and traditional elements restoration), desired demands (e.g., interactivity, expressiveness and modern aesthetic combination) and charming demands (e.g., innovativeness, personalized design and unique expression of cultural symbols) of the virtual image are clarified and the coefficients of satisfaction for each demand are calculated for sorting and weight division. On this basis, the design of the virtual image of Gannan Tea Picking Opera is carried out.

Based on the KANO model, this paper formulates the overall program of virtual image design and conducts user testing and feedback collection. The test results show that the designed virtual image performs well in terms of user satisfaction and acceptance. Most users believe that the virtual image retains the traditional flavor of the Gannan Tea Picking Opera, but is also modern and interactive, which can effectively stimulate their interest in this traditional form of opera. Especially among the young user group, the virtual image has significantly increased the popularity and influence of the tea picking opera through the dissemination of social media and short video platforms.

Keywords: Gannan tea picking opera · Virtual image design · Kano model · Traditional opera

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 37–47, 2025.
https://doi.org/10.1007/978-3-031-94171-9_4

1 Introduction

As a national intangible cultural heritage, Gannan Tea Casting Opera carries rich regional cultural connotations, but its dissemination and inheritance in the modern society faces challenges. The rise of digital technology provides new opportunities for the innovative dissemination of traditional culture, and virtual image design becomes an important way to promote the culture of Gannan Caicha Opera. Through virtual image design, the cultural elements of Gannan Tea Picking Opera can be combined with modern technology to create more attractive and contemporary cultural products, so as to expand its audience and enhance its cultural influence. And Kano model, as a scientific user demand analysis tool, can help designers accurately locate user demand, optimize design solutions and enhance user satisfaction by dividing demand into five attributes: basic, expectation, charismatic, irrelevant and reverse. The model has been successfully applied in the field of cultural and creative product design, but the research in the field of virtual image design is still in the exploration stage.

This study introduces the Kano model, which aims to systematically analyze the virtual image design of the Gannan Tea Picking Opera, clarify the classification and priority of user needs, formulate targeted design strategies, and optimize user experience. This not only helps to identify user needs scientifically, but also provides theoretical support and practical guidance for the digital inheritance and innovation of Gannan Tea Casting Opera, and promotes its dissemination and development in the new era.

2 Theoretical Foundations and Related Concepts

2.1 Gannan Tea-Picking Opera

Gannan Tea Picking Opera is a unique local opera in Gannan, with a long history and rich cultural connotation. Its origin can be traced back to the area of Anyuan and Xinfeng in Gannan, and it was initially a "tea picking song" for tea farmers, which was gradually combined with dance over time and developed into "tea picking lanterns" performed during the Lantern Festival, and gradually evolved into a mature form of opera during the Qing Dynasty [1]. In the Qing Dynasty, it gradually evolved into a mature form of opera [1]. Gannan tea-picking opera has become an important part of Hakka culture because of its profound regional characteristics and unique artistic style.

Gannan Tea Casting Opera has a unique form of artistic expression, lively and witty performance style, vivid and popular language, close to folk life, and its performance form is based on the structure of "triangular class" with "two dans and one clown", with distinctive characters, good at showing farming life, folk customs and love stories, which is popular among the audience [1]. They are good at showing farming life, folk customs and love stories, and are popular among the audience [2]. In terms of music, Gannan Tea Picking Opera is based on Gannan folk songs, and has gradually formed a unique vocal system, such as "Tea Cavity" and "Lamp Cavity", which have smooth melodies and cheerful rhythms, and are very much in tune with the local flavor [3]. In addition, most of its performance movements come from the body language of tea plucking, and the dance movements are light and flexible, fully demonstrating the life of the Hakka people.

In recent years, with the advancement of intangible cultural heritage protection, the tea picking opera has gradually received attention, but it also faces the impact of modernization, especially the loss of audience and the problem of a single mode of dissemination. For this reason, the researcher suggests empowering cultural communication through digital technology, using virtual images to innovate cultural communication methods, and enhancing the sense of identity of young groups.

2.2 KANO Model

Kano model, proposed by Noriaki Kano in 1984, is an analytical tool based on the relationship between user satisfaction and demand attributes. The model classifies user needs into five categories: basic needs, desired needs, charismatic needs, irrelevant needs and reverse needs, which can help designers clarify the priority of needs and optimize the design process. Academics have long conducted many researches and applications on kano model. Tang Zhongjun et al. combined the Kano model and fuzzy clustering method to propose a personalized demand acquisition method applicable to mass customization production, and successfully verified the practicality of the model through cases [4]. Tu Haili et al. applied the Kano model to WeChat's functional requirements analysis and found that users had higher expectations for core functions, while showing less demand for certain functions [5]. Wang Frost et al. constructed a user demand indicator system based on the Kano model and optimized the design strategy of intelligent products by analyzing the multi-dimensional analysis [6].

The application of Kano model in virtual image design not only helps to analyze and classify user needs, but also can effectively optimize the design process. Through the guidance of this model, the virtual image design of the Gannan Tea Picking Opera can more accurately refine the cultural elements, optimize the user interaction design, and provide scientific theoretical support and practical solutions to improve user satisfaction. At the same time, the KANO model prioritizes the requirements by the type of requirements and the Better-Worse coefficient, which accurately defines the development direction of the virtual image of Gannan Hakka Tea Casting Opera.

2.3 Virtual Image Design

Virtual image design is an interdisciplinary field integrating artistic creativity and cutting-edge technology, which has attracted much attention in recent years with the rapid development of digital technology. It uses computer graphics, animation technology and artificial intelligence to create virtual characters with unique appearances, personalities and behavioral patterns, which are not only widely used in traditional entertainment industries such as games, film, television and animation, but also show great potential in emerging fields such as live broadcasting, social media, education and medical care.

From the design point of view, virtual image design covers a variety of aspects such as character modeling, color matching and costume design. Designers need to have solid artistic skills and a keen insight into pop culture in order to create virtual characters that are aesthetically pleasing to users as well as conveying specific emotions and values. With technological advances, the interactivity of avatars is increasing, for example, through motion capture, voice synthesis and artificial intelligence algorithms,

which enable avatars to interact with users in real time and show more natural and vivid behavioral expressions. Meanwhile, academics have conducted a series of studies on virtual image design. Niu Min et al. pointed out that virtual image plays an important role in brand image building by virtue of its unique style and powerful communication ability. Virtual image can establish emotional connection with consumers, thus enhancing the market competitiveness of the brand [7]. Virtual images also show great potential in the protection and promotion of intangible cultural heritage. For example, Tan Xiaofeng et al. explored how to utilize 5G technology to create a virtual image of the non-genetic inheritors of the Jing ethnic group. By extracting the cultural genes of the Jing ethnic group and combining them with virtual interactive experiences, the study successfully realized the cross-temporal dissemination and rejuvenation promotion of ICH [8]. Drawing on these successful experiences will help enhance the attractiveness and cultural dissemination effect of the virtual image of the Gannan Tea Picking Opera, especially among young audience groups.

In this study, the virtual image, as a combination of digital technology and traditional culture, opens up a new path for the inheritance and promotion of Gannan tea-picking opera. It vividly shows the culture of tea-picking opera through costumes, actions and language, breaking through time and space limitations and attracting more audiences, including young groups. Its interactive function supports voice recognition and expression interaction, enhancing the audience's sense of immersion and participation. At the same time, the virtual image incorporates modern design elements, such as exquisite styling and fashionable clothing, which not only preserves the cultural value of traditional opera, but also fits modern aesthetics and enhances user experience. The realization of these functions provides strong support and new ideas for the dissemination of Gannan tea picking opera and the protection and inheritance of intangible cultural heritage.

3 User Requirements Analysis and Model Construction

3.1 Indicator Construction and Questionnaire Design

In constructing the demand indicator system, this study combines fieldwork and literature research to ensure that the study is both theoretically grounded and in line with actual needs. Through in-depth research on the traditional elements of Gannan Hakka Tea Picking Opera and extensive literature review, the research team systematically sorted out the key elements of virtual image design and extracted the core keywords of "Gannan Hakka Tea Picking Opera Virtual Image Design" by using the KJ method. After rigorous checking and organizing, 13 core demand function points were finally identified (see Table 1), which provided clear guidance for the questionnaire design.

In order to comprehensively assess users' needs and expectations for virtual image design, this study designed a demand questionnaire based on the KANO model. The questionnaire is designed with both positive and negative questions for each demand point, for example, for demand point D1, the positive question "How do you feel about the reduction of traditional clothing style by avatars?" and the negative question "How do you feel about the reduction of traditional clothing style by avatars?" For example, for point D1, the positive question "What do you think about the virtual image's reproduction of traditional clothing styles?" and the negative question "What do you think about the

virtual image's failure to reproduce traditional clothing styles?" This kind of positive and negative design can be more accurate. This positive and negative design can more accurately capture users' real feelings and demand priorities. The questionnaire is rated on a 5-point Likert scale, with options such as "I like it", "I take it for granted", "I don't care", "I can't accept it", "I can't accept it", "I can't accept it", "I can't accept it", and "I like it". " "Unacceptable", which effectively quantifies the subjective evaluation of users (see Table 2). After the questionnaire was completed, the research team distributed it widely through multiple channels to collect enough sample data to provide strong support for subsequent analysis and decision-making.

Table 1. Attributes of virtual image design for Gannan Hakka tea-picking opera.

D1	Restoration of traditional dress styles	D10	Adaptation of virtual images
D2	Traditional make-up and headdress	D11	Fusion of modern art elements
D3	Realistic facial expressions	D12	Precise reproduction of dance movements
D4	Preservation of classic props for tea-picking theatre	D13	Gender diversity in virtual images
D5	Personalisation and creative expression of cultural symbols		
D6	Modernisation of clothing and make-up		
D7	Modern aesthetic elements of colour schemes		
D8	Variety of styling options		
D9	Interactivity and dynamic change		

Table 2. Kano model classification evaluation table.

User Needs	Likes	Deserved	Doesn't matter	Barely accepts	Dislikes
Like	Suspicious outcome	A	A	A	O
Deserved	R	I	I	I	M
Doesn't matter	R	I	I	I	M
Barely accepts	R	I	I	I	M
Dislikes	R	R	R	R	Suspicious outcome

Table 3. Kano model classification of design attributes of Gannan Hakka tea-picking opera virtual image.

Causality	A/%	0/%	M/%	I/%	R/%	B(Better)	W(Worse)	Typology
D1	18.34	20.55	56.66	4.45	0.00	0.39	−0.78	M
D2	17.77	27.78	50.01	4.44	0.00	0.46	−0.78	M
D3	25.00	18.33	48.33	8.34	0.00	0.43	−0.67	M
D4	20.00	41.11	29.44	9.45	0.00	0.61	−0.71	O
D5	37.78	22.22	32.22	7.78	0.00	0.6	−0.55	A
D6	42.78	29.44	23.89	3.89	0.00	0.72	−0.53	A
D7	23.89	48.89	16.11	11.11	0.00	0.73	−0.66	O
D8	27.78	35.00	25.56	11.66	0.00	0.63	−0.61	O
D9	30.56	41.11	21.11	7.22	0.00	0.72	−0.62	O
D10	12.78	45.00	33.89	8.33	0.00	0.58	−0.79	O
D11	18.89	42.78	28.89	9.44	0.00	0.62	−0.72	O
D12	33.33	27.78	25.00	13.89	0.00	0.61	−0.53	A
D13	17.22	38.89	31.67	12.22	0.00	0.56	−0.71	O

3.2 Data Collection and Processing

The questionnaires were distributed online and offline, targeting design experts, related industry workers and some college students, with a two-week cycle, and a total of 200 questionnaires were sent out, of which 27% were from design experts, 55% were from college students, and 18% were from related industry workers. After cleaning the questionnaire data and eliminating the short answer time and invalid answers, 180 valid questionnaires were finally screened out. The collected valid questionnaire data are statistically categorized according to the KANO model evaluation results classification table, where A represents charismatic attributes, O represents expected attributes, M represents necessary attributes, and I represents non-differentiated attributes, and the specifics are shown in Table 3. Based on the above analysis results, the coefficient of influence of satisfaction (Better) and the coefficient of influence of dissatisfaction (Worse) of each index are calculated, and the value of the proportion of attributes is substituted and calculated. The value of the proportion of each attribute is substituted and the satisfactory influence value is obtained, as shown in the following formula for calculating the satisfactory influence coefficient B.

$$B = (A + O)/(A + O + M + I)$$

The larger value resulting from this calculation indicates a greater impact on user satisfaction. Substituting the values of the percentage of each attribute and finding the dissatisfaction influence value is shown in the following formula for the dissatisfaction influence coefficient W.

$$W = -1 \times (O + M)/(A + O + M + I)$$

The smaller the value resulting from this calculation indicates a greater impact on user dissatisfaction.

3.3 Classification and Sorting of Demand Categories

The satisfaction index of each previously obtained demand is sorted out, see Table 3, and the priority of all functional demands of Gannan Hakka Tea Picking Opera virtual image design is analyzed and calculated, and sorted and organized, the specific results are shown in Table 4.

Table 4. Prioritization of the importance of the virtual image needs of Gannan tea picking opera.

Typology	Prioritisation of all functional requirements
M	D2 > D1 > D3
O	D7 > D10 > D9 > D11 > D4 > D13 > D8
A	D6 > D5 > D12

(D1); the desired attributes to be focused on are "modern aesthetic elements of color matching" (D7), "adaptability of virtual image" (D10), and "interactivity and dynamic change of virtual image" (D11). Interactivity and Dynamic Change" (D9), and because of the large number of desired attributes of the avatars, the development of those that are in the lower order or are more difficult to develop can be put on hold, such as 'Variety of Styling Options' (D8) and 'Gender Diversity of Avatars' (D9). "gender diversity of avatars" (D13); finally, for the charm attributes, 'modernization of dress and makeup' (D6) and 'personalization and innovative expression of cultural symbols' (D5) are also desired attributes. Finally, "modernization of clothing and makeup" (D6) and "personalization and creative expression of cultural symbols" (D5) are also considered.

4 Design Program and Implementation Process

4.1 Design Strategy and Process

In the development of the virtual image design of Gannan Hakka Tea Picking Opera, this study adopts a systematic design strategy, comprehensively utilizing the literature review, field research and the data of the previous research to ensure that the design scheme is both scientific and in line with the actual needs. Through these methods, the cultural connotation of Gannan Hakka Tea Picking Opera and the aesthetic needs of modern audiences are comprehensively grasped to lay a solid foundation for the virtual image design. The specific process is as follows: firstly, the demand elements after rigorous analysis and screening are introduced into the design process to generate the preliminary visual framework of the virtual image; subsequently, the multi-dimensional presentation capability of the virtual image is further optimized to achieve dynamic interaction and multi-angle display in the virtual environment, so as to enhance its expressive power and user experience.

4.2 Achievement Display and Effect Demonstration

According to the design process, the elements that need to be prioritized through the Kano model and the satisfaction coefficient calculation in Chapter 4 are sorted out, and the following needs are finally concluded to be satisfied through the styling design of the virtual image: "traditional make-up and headdress", "restoration of traditional clothing style", "modern aesthetic elements of color matching", "modern adjustment of dress and makeup", "personalization and innovative expression of cultural symbols". As a result, the visual image of the virtual image as well as clothing and accessories, etc. is designed by combining these design elements.

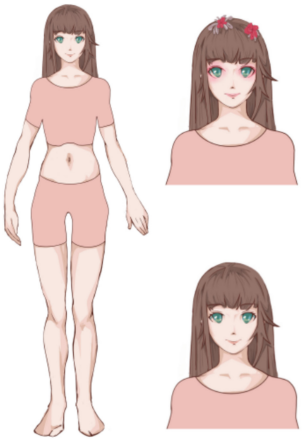


Fig. 1. Virtual image modeling design of Gannan tea picking opera.

The virtual image is named "Cai Ling", which combines the elements of "tea picking" and "spirituality", implying that the character is free and full of vitality on stage. In terms of styling, the main body refers to the height and body shape of an adult woman with fair skin. For the facial design, teal-colored hair and green eyes are used, which further emphasize the theme of traditional Chinese elements. In terms of accessories, the traditional silver headdress and peony flowers of the Gannan Hakka people were added, and the makeup of the tea picking opera was matched to the virtual image in the specific styling, which is shown in Fig. 1.

In terms of costume design, two sets of costumes were designed to meet the diversity and personalization needs of the modeling. One set is rich in elements of Gannan Hakka Tea Picking Opera, visually presenting the theme of the virtual image; the other set is in a casual style, incorporating elements of Tea Picking Opera only in the color scheme and some details, the specific design is shown in Fig. 2. Applying the virtual image to other carriers can broaden the publicity channels of Gannan Tea Picking Opera and attract more attention to the local culture.

The virtual image can be applied to the field of live broadcasting, transforming it into a virtual anchor. Through specialized software for motion capture and sound acquisition, the virtual image "Cai Ling" can interact with the audience in real time (see Fig. 3). This

Fig. 2. Costume design of virtual image of Gannan tea picking opera.

Fig. 3. Gannan Tea Picking Opera Virtual Image Live Showcase.

Fig. 4. Application of virtual image of Gannan tea picking opera.

approach not only meets the interactivity and dynamic change requirements of the virtual image, but also better promotes the cultural dissemination of Gannan tea picking opera. In addition, in order to meet the adaptability requirements of the virtual image, it can be

presented on a variety of devices. For example, real-time live broadcasting can be carried out on mobile devices such as cell phones and tablets, while promotional introduction videos can be played in the display screens of shopping malls (see Fig. 4). By these means, the channels of cultural dissemination are further broadened.

4.3 Feedback Collection

In order to assess the acceptance and effectiveness of the designed virtual image of Gannan Hakka Tea Picking Opera, the research team extensively collected public opinions through questionnaires. The questionnaire consisted of 10 questions, involving information such as satisfaction with the appearance of the virtual image, the degree of restoration of the elements of the tea picking opera, whether or not they were willing to watch the live performance of the virtual image, and the age of the survey respondents. A total of 100 questionnaires were distributed, and 92 valid questionnaires were finally screened out after data cleaning and elimination of invalid responses.

The results of the survey showed that 62% of the respondents gave positive comments on virtual images. It is worth noting that respondents aged between 10 and 30 years old were as high as 72% satisfied with avatars, which is significantly higher than the 53% of respondents over 30 years old. This result suggests that the avatar has a high level of attraction among the younger group.

In summary, most users believe that the virtual image can effectively represent the Gannan Tea Picking Opera and play a positive role in publicity, enhance the popularity and influence of the opera, and attract more people, especially the young group, to pay attention to the opera itself. Therefore, the design program has high feasibility and practical application value.

5 Summary and Reflection

Taking Gannan Tea Picking Opera, a national intangible cultural heritage, as the research object, this paper explores the new path of digital communication of traditional culture through the modernized means of virtual image design, providing new ideas and practical methods for cultural inheritance. The study deeply analyzed the historical background, artistic characteristics and cultural connotation of the Gannan Tea Picking Opera, refined the core cultural elements required for virtual image design, and introduced the Kano model to classify and analyze the users' needs, clarified the users' basic needs, desired needs and charismatic needs, and carried out the design in this way.

Although certain research results have been achieved, this paper still has limitations. First, the sample coverage of the study is limited, mainly focusing on user groups with specific cultural backgrounds and age levels, and future research needs to expand the sample scope to improve the generalizability of the results. Second, the research method is relatively single, and a combination of quantitative methods can be used in the future to enhance the objectivity and credibility of the study. In addition, due to technical limitations, the interactive function and dynamic performance of avatars still need to be improved, especially in terms of intelligent interaction and emotion recognition, which require more advanced technical support.

In summary, this paper analyzes the user demand through Kano model, provides a scientific analysis framework and implementation path for the virtual image design of Gannan Tea Picking Opera, and verifies the effectiveness and innovativeness of the virtual image design in cultural inheritance. The research results not only provide a new methodological reference for the digital communication of traditional culture, but also provide a reference for the digital protection and promotion of other intangible cultural heritage. The future research will continue to explore the combination of virtual technology and cultural innovation to inject more vitality and possibilities into the modernization and development of intangible cultural heritage.

References

1. Wei, X.: Cultural and creative product design of Gannan tea-picking opera. Packaging Eng. **41**(20), 246–249+258 (2020). https://doi.org/10.19554/j.cnki.1001-3563.2020.20.040
2. Zhong, J.: On the aesthetic characteristics of Gannan tea-picking opera. Jiangxi Soc. Sci. (4), 197–200 (2008)
3. Su, Z.: On the artistic form and innovation of Gannan tea-picking opera. Hundred Schools Arts (2), 9–12 (2003)
4. Tang, Z., Long, Y.: Research on personalized demand acquisition method based on kano model. Soft Sci. **26**(2), 127–131 (2012)
5. Tu, H., Tang, X.: KANO model analysis of Wechat functional requirements. J. Intell. **34**(5), 174–179 (2015)
6. Wang, S., Yin, G., He, Z.: Research on user demand indicator system based on kano model. Packaging Eng. (4), 209–210+213 (2006)
7. Niu, M., Chen, G.: Brand image design and communication based on virtual idol symbols. Packaging Eng. **42**(18), 282–286 (2021). https://doi.org/10.19554/j.cnki.1001-3563.2021.18.034
8. Tan, X.: Research on virtual image design and promotion of Guangxi Jing ethnic group's intangible cultural heritage inheritors in the 5G Era. Packaging Eng. **45**(2), 480–490 (2024). https://doi.org/10.19554/j.cnki.1001-3563.2024.02.053

Exploring User Experiences in Hybrid Cultural Visits: Insights from a User Study

Zhuosi Lai(✉), Yi Zhao, Xin Cheng, Rodrigo Hernández-Ramírez, and Luke Hespanhol

The University of Sydney, School of Architecture, Design and Planning, Camperdown, NSW 2050, Australia
{zhuosi.lai,yi.zhao1,dylan.cheng, rodrigo.hernandez,luke.hespanhol}@sydney.edu.au

Abstract. Traditional on-site and remote cultural visits face challenges such as geographical constraints, limited accessibility, and difficulties in promoting presence and social engagement for remote visitors. Hybrid cultural visits, integrating physical and virtual elements, offer a promising solution to enhance accessibility and inclusivity while combining the strengths of both formats. To explore future design directions, we conducted a usability study to evaluate user experiences during hybrid gallery visits, focusing on communication, engagement, and emotional connectivity. The study employed a two-round methodology: one participant engaged with the gallery while interacting with a researcher via video, and vice versa. Quantitative data, collected through questionnaires, assessed mental immersion and social richness. The findings revealed that remote participants (RPs) experienced diminished control, passive engagement, and weaker emotional connections with both exhibits and on-site participants (OPs) compared to their on-site counterparts. These results highlight the limitations of current hybrid models for creating immersive and interactive remote experiences. The study contributes to research on hybrid cultural spaces, suggesting avenues for future studies and more inclusive frameworks for the Galleries, Libraries, Archives, and Museums (GLAM) sector.

Keywords: GLAM Sector · Hybrid Visiting · Immersive Experience

1 Introduction

Cultural institutions, such as galleries, libraries, archives, and museums (GLAM), play a vital role in preserving and disseminating cultural heritage. Traditionally, these spaces have relied on physical visits, allowing visitors to engage directly with exhibits and artifacts. However, such visits are often constrained by geographical barriers, limited accessibility for individuals with disabilities, and capacity restrictions [10]. In recent years, the rise of digital technologies has enabled online gallery visits, which offer greater reach and flexibility. Yet, these

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 48–57, 2025.
https://doi.org/10.1007/978-3-031-94171-9_5

virtual experiences often fall short in fostering a sense of presence, emotional connection, and social engagement, particularly for remote participants [7].

To address these limitations, hybrid cultural visits have emerged as an innovative approach that integrates physical and virtual elements. By combining the strengths of on-site and online experiences, hybrid visits aim to enhance accessibility, inclusivity, and interactivity for a broader audience. However, despite their potential, current common hybrid approaches—such as the use of teleconference apps on smartphones—often fail to provide RPs with the same level of engagement, control, and emotional connectivity as on-site visits. This gap highlights the need for further exploration into the design and implementation of hybrid cultural experiences.

In this study, we analysed the user experience of hybrid gallery visits, focusing on communication, engagement, and emotional connectivity between on-site participants (OPs) and remote participants (RPs). Through a two-round approach that involved 27 participants, we evaluated experiential differences between these groups, using structured questionnaires to assess mental immersion and social richness [1] [8]. Our findings revealed significant disparities in control, engagement, and emotional connection, underscoring the challenges faced by RPs in commonly adopted approaches to hybrid experiences. These insights contribute to the growing interest in hybrid cultural spaces, offering valuable implications for the design of more inclusive and engaging frameworks in the GLAM sector.

2 Usability Study

2.1 User Study

To gather information on current hybrid visit experiences and identify potential design opportunities, we conducted a field study evaluating the experience of hybrid visits in a gallery setting, adopting a setup understood as familiar and intuitive to all participants. Namely, we established the hybrid visit as a call via a widely used video communication platform, Zoom[1]: on-site visitors had the platform active on their smartphones during their in-person visit to the gallery, while remote visitors joined the call from wherever they were, via their personal devices. We hoped that this basic setup would provide a foundational understanding of the user experience associated with hybrid visits and generate insights into how digital technologies can be redesigned to enhance it.

The study followed two rounds:

– Round 1: On-site Study. In the first round, participants were physically present in the gallery, while researchers joined remotely via Zoom. The OP explored the gallery space and interacted with the researcher through the Zoom application installed on their phone.

[1] https://www.zoom.com.

– Round 2: Remote Study. In the second round, the roles were swapped, with participants joining remotely while the researcher was on-site at the gallery. The RP used Zoom to view the gallery exhibits and interact with the on-site researcher.

Given the required logistics and to allow for geographically distant participants, we adopted a between-subject study design, i.e., each participant only experiences a single scenario. However, participants were allowed to choose whether they wanted to participate as an on-site or remote visitor. To ensure the experience closely replicated a natural gallery visit, no predefined script was provided for the interactions. Participants were encouraged to engage freely in discussions about the exhibits with the researcher, mirroring the typical dialogue that might occur during a gallery visit. Participants were also asked to make spontaneous comments and observations about the exhibits as they explored the venue.

2.2 Participants

We recruited 27 participants (14 women, 13 men) between the ages of 18 and 49. Participants were recruited through flyers posted on campus and shared through online social media channels. All participants voluntarily participated in the study after making initial contact in response to recruitment materials. Before the study, participants received a participant information statement and a consent form to view and sign. The participants were divided into two groups according to their preferences: 13 OPs and 14 RPs. The study was carried out according to ethical guidelines and the research protocol was reviewed and approved by the Human Research Ethics Committee of the University of Sydney.

2.3 Study Procedure

The OPs (Fig. 1) were instructed to act as tourists, exploring the gallery in person. Their role was to engage with the exhibits and provide feedback as they would during a typical gallery visit. OPs received a mobile device with Zoom to facilitate communication with their remote partner (acted by a researcher). This method allowed them to share their view of the exhibits in real time and also enabled the remote researcher to see and interact with the gallery space through the OP's perspective.

The session was accessed using a computer or mobile device via Zoom and communicated with the partner on site (acted by a researcher) throughout the visit. RPs ask questions, make observations, and explore the gallery exhibits with their on-site partners.

The shared tour session typically lasted between 10 to 20 min. After the hybrid visit, all participants were asked to complete a post-study questionnaire to collect data on their experiences, focusing on the effectiveness of the hybrid interaction, the level of engagement, and their perception of presence during the tour.

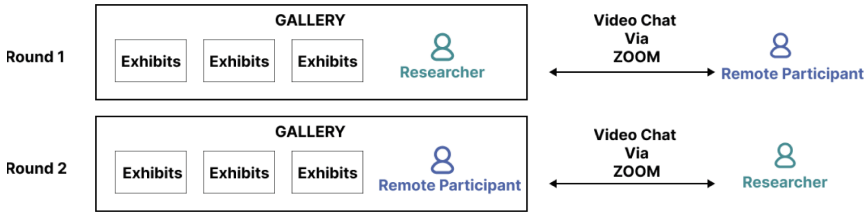


Fig. 1. Details of User Study Progress.

2.4 Data Collection

The study collected quantitative data using a combination of questions from two structured questionnaires. The questions were adapted from the Networked Minds Social Presence Inventory [1] to measure emotional contagion, communication clarity, and behavioural interdependence, and from the Temple Presence Inventory (TPI) [8] to assess mental immersion and social richness. The questions were grouped into four key categories: 1) Participants' Perception of Self (Q1-Q3); 2) Participants' Perception of Others (Q4-Q6); 3) Mental Immersion (Q7-Q11) and 4) Social Presence (Q12-Q18). The questions are listed in Table 1.

Table 1. Questions in the Study Categorised by Participants' Perception, Mental Immersion, and Social Presence.

Participants' Perception of Self (Q1-Q3)	
Q1	How much were you influenced by your visiting partner's moods?
Q2	Were you able to communicate your thoughts and intentions clearly to your visiting partner?
Q3	How much was your behaviour affected by your visiting partner's behaviour or remarks?
Participants' Perception of Others (Q4-Q6)	
Q4	How much was your visiting partner influenced by your moods?
Q5	Was your visiting partner able to communicate their thoughts and intentions clearly to you?
Q6	How much was your visiting partner's behaviour affected by your own behaviour or remarks?
Mental Immersion (Q7-Q11)	
Q7	To what extent did you feel mentally immersed in the joint visit experience?
Q8	How involving was the experience?
Q9	How completely were your senses engaged?
Q10	To what extent did you experience a sensation of visiting the museum/gallery together with your visiting partner?
Q11	How relaxing or exciting was the experience?
Social Presence (Q12-Q18)	
Q12	Remote/Immediate—Please choose the number that best describes your evaluation of the visiting experience.
Q13	Unemotional/Emotional—Please choose the number that best describes your evaluation of the visiting experience (Unemotional/Emotional).
Q14	Unresponsive/Responsive—Please choose the number that best describes your evaluation of the visiting experience.
Q15	Dead/Lively—Please choose the number that best describes your evaluation of the visiting experience.
Q16	Impersonal/Personal—Please choose the number that best describes your evaluation of the visiting experience.
Q17	Insensitive/Sensitive—Please choose the number that best describes your evaluation of the visiting experience.
Q18	Unsociable/Sociable—Please choose the number that best describes your evaluation of the visiting experience.

2.5 Data Analysis

We first assessed the internal consistency of the questionnaires completed by both OP and RP using Cronbach's Alpha [4] [11]. The results showed excellent internal reliability, with a Cronbach's Alpha of $\alpha = 0.93$ for both OP and RP groups, confirming that the questionnaires were highly consistent and suitable for further analysis.

Next, we performed a descriptive analysis of the data to obtain an overview of the participants' responses in both groups (OP and RP). For each variable, we calculated the means μ and standard deviations σ to explore the central tendency and variability of the data. Following descriptive analysis, we examined the normality of the data using Shapiro-Wilk tests for each variable. The results ($p > 0.05$) confirmed that the data were normally distributed across all variables, allowing us to proceed with parametric tests. Based on the results of the normality tests, we performed an independent samples T-test to compare the means of the two groups (OP and RP) for each variable. Before performing the T-tests, we used the Levene test to determine whether the assumption of equal variances was valid. If the variances were equal ($p > 0.05$), we used the "Equal variances assumed" T-test results; otherwise, we applied the "Equal variances not assumed" approach.

3 Results

For the category of participant self-perception, which includes Q1 (Mood), Q2 (Communication) and Q3 (Behaviour), the results of the T-test show no significant differences between OP and RP. This suggests that the self-perception of the participants remained stable, regardless of whether they were physically present or participating remotely (see Table 2 and Fig. 2).

Moving to participants' perception of others (Table 2), while there was no significant difference for Q4 (Mood) and Q6 (Behaviour), the T-test results for Q5 (Communication) show a significant difference ($t = -2.207, p = 0.037$). This shows that RPs felt they were better able to understand and respond to the communication from their on-site partners compared to OPs. The results suggest a dynamic where RPs were more focused on interpreting the communication of OPs, potentially due to the unidirectional nature of the experience, where RPs are more dependent on the OPs for guidance. Although the results for Q6 (Behaviour) did not reach significance, the trend suggests that RPs felt less able to influence their partners' behaviour, reinforcing the expectation that the hybrid experience is more passive for the RPs (Fig. 2).

The results of mental immersion do not show statistically significant differences between OPs and RPs (Table 2). However, trends in the data indicate that RPs generally reported lower levels of engagement and immersion. For example, while the p-value for Q9 (Immersion) was not significant, the data suggests a reduction in immersion scores for RPs compared to OPs. This is due to the inherent limitations that RPs experienced, which restrict participants to auditory and visual stimulation, without the full sensory experience available to those

Table 2. T-Test Results Comparing OP and RP Across All Study Questions, statistically significant results are marked with an asterisk (*), showing $p < 0.05$.

Question	OP		RP		t-value	p-value
	Mean	SD	Mean	SD		
Q1 Mood	5.00	1.00	4.79	1.48	0.44	0.665
Q2 Communication	5.46	1.27	5.57	1.52	-0.20	0.841
Q3 Behaviour	4.54	0.88	4.14	1.91	-1.68	0.501
Q4 Mood (Other)	4.31	1.32	3.59	1.55	1.29	0.210
Q5 Communication	4.85	1.21	5.85	1.15	-2.21	**0.037***
Q6 Behaviour (Other)	4.77	0.83	4.12	1.38	1.46	0.157
Q7 Visit Experience	5.08	1.26	4.54	1.73	0.50	0.622
Q8 Engagement	5.38	1.30	5.00	1.40	0.71	0.959
Q9 Immersion	5.23	1.36	4.62	1.68	1.03	0.313
Q10 Companionship	5.31	1.12	4.85	1.38	1.18	0.366
Q11 Stimulating	5.15	1.47	4.77	1.44	0.74	0.843
Q12 Presence	5.08	1.39	4.87	1.51	0.50	0.620
Q13 Affectivity	5.15	0.90	3.91	1.64	2.42	**0.023***
Q14 Interactivity	5.38	1.09	4.69	1.53	1.43	0.612
Q15 Vibrancy	5.15	1.38	4.54	1.66	1.05	0.737
Q16 Intimacy	4.92	1.22	4.23	1.66	1.36	0.823
Q17 Sensitivity	5.31	1.21	4.62	1.52	1.26	0.169
Q18 Gregariousness	5.23	1.36	4.77	1.48	0.76	0.509

physically present. Similarly, RPs reported feeling less connected to their on-site partners (Q10). These findings highlight a gap in the depth of mental immersion experienced by RPs, pointing to areas where technological enhancements could foster a more immersive remote experience.

In the category of social presence, the T-test results for Q13 (Affectivity) revealed a significant difference ($t = 2.418, p = 0.023$), with OPs reporting higher levels of affective connection than RPs. This finding underscores the emotional disconnect that RPs often experience in hybrid settings, where the physical absence diminishes the sense of emotional engagement. Although the results for Q12 (Presence) were not significant, the data shows a similar trend, with RPs reporting a weaker sense of co-presence compared to OPs. This pattern is consistent with other social presence variables, suggesting that RPs feel less emotionally and socially connected to OPs.

Overall, these results, guided by the t-test analysis, demonstrate that while both OPs and RPs had generally positive experiences, RPs faced challenges in terms of communication, emotional connection, and immersive engagement. These findings highlight the limitations of current remote interaction technologies, particularly in fostering a sense of presence and emotional involvement

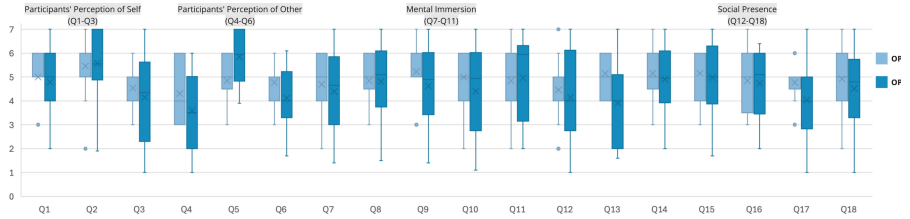


Fig. 2. Results of the questionnaire comparing the distribution of responses from OP and RP across four categories. Each box shows the interquartile range (IQR), with the horizontal line representing the median, with 'X' = μ. The whiskers represent the full range of non-outlier data, while the dots indicate outliers.

for RPs. Significant differences in communication and emotion, combined with trends in mental immersion and social presence, not only suggest that improvements in the design of hybrid visiting technologies are needed to solve the experiential gap between OP and RP, but point to which experiential aspects to focus on.

4 Discussion

The findings of this study show both the transformative potential and the existing limitations of hybrid cultural visits in creating inclusive and engaging experiences for on-site and remote visitors. Hybrid visits can offer a promising solution to the suboptimal engagement of current online visits, as many individuals lack the digital literacy or cultural capital to effectively navigate museum websites [9]. Real-time conversations during these visits help foster deeper engagement and a more immersive experience. Although RPs reported challenges in control, emotional connection, and social presence during the study, the hybrid mode still demonstrated its ability to expand and unlock access to cultural spaces for individuals who cannot physically visit galleries due to geographical, physical, or other constraints. This dual perspective underscores the importance of refining hybrid visit designs to maximise their benefits while addressing current shortcomings, ensuring that cultural spaces become more accessible, inclusive, and engaging for all.

4.1 Expanding Access to Cultural Spaces

One of the most compelling advantages of hybrid visits is their ability to increase accessibility to cultural spaces for individuals who cannot visit in person due to distance, physical or psychological disabilities, financial constraints, or health concerns. While fully remote visits, facilitated by tools like on-site robots or VR technology, also enhance accessibility, hybrid visits uniquely integrate the irreplaceable sensory and emotional experiences of on-site visitors with the accessibility of remote participation [7]. This integration ensures that cultural spaces are

no longer confined by geographical or physical boundaries. For instance, someone living in an area without galleries can now experience cultural programs in remote venues that were previously inaccessible. By bridging physical and virtual participation, hybrid visits create a more inclusive and engaging model for cultural engagement. With minimal need for additional facilities, hybrid visits enable cultural institutions to reach a wider and more diverse audience, aligning with the GLAM sector's growing emphasis on inclusivity [4], which is not just a technological achievement but also a social and cultural achievement.

4.2 Challenges for Remote Participants

Despite their potential, the current hybrid model presents some challenges, especially for remote visitors. While hybrid visits can improve accessibility, our study revealed disparities in the quality of experience between OPs and RPs. According to the study result, remote visitors can experience reduced control, passive participation, and weaker emotional connections compared to on-site visitors, raising critical questions about how to expand access without compromising fairness in participation. These challenges are partly due to the unidirectional nature of the experience, where remote visitors rely heavily on on-site visitors for guidance, limiting their engagement. A key issue here is the weaker emotional connection, as reported by RPs during the study: emotional engagement, a cornerstone of cultural experiences, plays important role for visitors to connect deeply with exhibits and with one another. The lack of physical proximity and sensory input, such as touch and spatial awareness limits the emotional richness of the experience for remote visitor [5], and as a result, remote visitors may feel disconnected from both the exhibits and the on-site partner. Additionally, studies on virtual participation in cultural settings have shown that remote visitors often feel excluded from the social fabric of shared experiences, leading to isolation and reduced satisfaction [2]. Emotional and social engagement are important elements transforming a simple gallery visit into a memorable and transformative experience, and lack of these can compromise the overall visiting experience.

To address these challenges, hybrid visits should enhance real-time interactions between remote and on-site visitors, fostering a sense of inclusion and shared experience. Research highlights the difficulties of promoting agency and emotional engagement in virtual environments [3], underscoring the need for innovative solutions. Future hybrid visit designs can focus on improving interactivity and empowering remote visitors to play a more active role in the experience: for example, incorporating tools that allow remote visitors to control camera angles, zoom in on exhibits, or even guide the on-site visitors could enhance their sense of immersion. Additionally, using immersive technologies like augmented reality (AR) or haptic feedback could provide RPs with a more emotionally engaging experience. Research by [6] suggests that interactive and immersive technologies can significantly enhance user engagement and emotional connection in virtual environments, where one of the findings highlight the potential of AR and haptic systems in creating a sense of presence and immediacy,

making remote visitors feel as though they are physically part of the experience, without it, RPs may feel detached from both the exhibits and the social dynamics of the visit, reducing the overall impact of the experience. It is notable that the integration of these technologies must be done thoughtfully, ensuring that they are accessible and intuitive for users with varying levels of digital literacy.

5 Conclusion

This study explored the user experience of hybrid gallery visits, with a focus on communication, engagement, and emotional connectivity between on-site and remote visitors. The findings reveal both the potential and the limitations of hybrid visits. On the one hand, hybrid visits offer a promising solution to expand access to cultural spaces for individuals who cannot physically visit galleries due to geographical, physical, or other constraints. However, remote visitors face challenges in control, emotional connection, and social presence, highlighting the need for technological and design innovations to create more equitable and engaging experiences.

The study contributes to the growing discourse on hybrid cultural spaces by identifying key challenges and opportunities for improvement. Specifically, it underscores the importance of improving interactivity, emotional engagement, and social presence for remote visitors. Future research may explore the use of immersive technologies (e.g. VR, AR) and bidirectional communication tools to bridge the experiential gap between on-site and remote visitors.

By addressing these challenges, cultural institutions can develop more inclusive and engaging frameworks for hybrid visits, ensuring that all participants, regardless of their physical location, can fully experience and connect with their cultural heritage. This study provides a foundation for future research and design efforts aimed at creating more equitable and meaningful hybrid cultural experiences.

References

1. Biocca, F., Harms, C.: Guide to the networked minds social presence inventory v. 1.2. Tech. rep., MIND Labs, Michigan State University (2002). https://web-archive.southampton.ac.uk/cogprints.org/6743/1/2002_guide_netminds_measure.pdf
2. Black, G.: Transforming Museums in the Twenty-First Century. Routledge, New York (2012)
3. Carrozzino, M., Bergamasco, M.: Beyond virtual museums: experiencing immersive virtual reality in real museums. J. Cult. Herit. **11**(4), 452–458 (2010). https://doi.org/10.1016/j.culher.2010.04.001
4. Cronbach, L.J.: Coefficient alpha and the internal structure of tests. psychometrika **16**(3), 297–334 (1951). https://doi.org/10.1007/BF02310555
5. Falk, J.H., Dierking, L.D.: The Museum Experience Revisited. Routledge, New York (2013)

6. Freeman, A., Becker, S.A., Cummins, M.: NMC/CoSN horizon report: 2017 k. Tech. rep, The New Media Consortium (2017)
7. Lai, Z., Hespanhol, L.: An investigation on strategies for remote interactions with cultural spaces. In: Proceedings of the 33rd Australian Conference on Human-Computer Interaction. p. 170-175. OzCHI '21, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3520495.3520523
8. Lombard, M., Ditton, T.B., Weinstein, L.: Measuring presence: the temple presence inventory. In: Proceedings of the 12th Annual International Workshop on Presence, pp. 1–15. International Society for Presence Research Los Angeles, CA (2009)
9. Mihelj, S., Leguina, A., Downey, J.: Culture is digital: cultural participation, diversity and the digital divide. New Media Soc. **21**(7), 1465–1485 (2019). https://doi.org/10.1177/1461444818822816
10. Pollalis, C., et al.: ARtLens: enhancing museum visitors' engagement with African art. In: Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems, pp. 195–200. DIS '18 Companion, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3197391.3205435
11. Taber, K.S.: The use of Cronbach's alpha when developing and reporting research instruments in science education. Res. Sci. Educ. **48**(6), 1273–1296 (2016). https://doi.org/10.1007/s11165-016-9602-2

Virtual Character Design Transformation of Dongba Cultural Information Based on AHP-Meme Theory

Anglu Li[1] and Chang Guo[2(✉)]

[1] Sichuan University, Chengdu 610207, Sichuan, China
[2] Tongji University, Shanghai 200092, China
guoagcchang@gmail.com

Abstract. This study focuses on the protection and inheritance of traditional national culture in online virtual character design. In the digital era, virtual characters have become an important visual means for individuals to express their ideas and cultural identity. The diversity of ethnic cultures is of great significance in enriching the culture of online communities and meeting the personalized needs of users. Dongba culture provides rich modeling and pattern materials for virtual image design with its unique visual elements. This study aims to analyze the development and characteristics of Dongba culture. And we establish a design method that combines Dongba culture with online virtual image desig to transform cultural information to design content. This study is based on the meme theory, which is derived from Darwin's theory of evolution, and regards meme as the unit of information in culture, like the genetic information in DNA. With meme theory, we extracted the "genetic" factors of Dongba culture and categorized them into layers, and then created a genetic map of Dongba culture. In addition, the importance of each cultural factor in the design of the virtual image was calculated with the AHP method, and their proportion in the design was determined. The conclusion shows that the cultural information transformed based on AHP-modeling theory can be effectively applied to virtual image design, providing designers with a convenient way to obtain design elements of Dongba culture. This comprehensive research method not only provides a reference for the development of national culture on digital media but also supports the cultural diversity and personalized needs of online virtual communities.

Keywords: Dongba culture · cultural genes · AHP · virtual character design

1 Introduction

With the development of digital media art and online community, Virtual character has become a part of people's online social platform [1], and it is a visual way for people to express their ideas, character, hobbies, culture and other information [2, 3]. The cultural elements of different regions and nationalities are the subjects that Virtual character design is keen to choose [3]. The long history of national culture can promote

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 58–75, 2025.
https://doi.org/10.1007/978-3-031-94171-9_6

the cultural diversity of the network virtual community and can also enhance the cultural demand and satisfaction of user groups of different cultural backgrounds on the Virtual character [4, 5]. At the same time, the protection and inheritance of traditional national culture is a long-term issue facing the development of human society [6]. The path of digital development of traditional ethnic cultures is one of the most important methods of passing on regional cultures today [7]. The Dongba culture of the Naxi people is the only culture that still uses hieroglyphics. On October 15, 2003, it was officially listed on the Memory of the World Register by UNESCO [8]. The rich visual elements of the Dongba culture can provide stylistic and textural references for Virtual characters [9]. However, there are three problems: how to deeply sort out the national symbols and stylistic characteristics contained in Dongba culture? How to analyze their respective importance in Virtual character design? How to help designers who do not understand the Dongba culture to quickly design design works containing the information of Dongba culture? This paper constructs a gene map of Dongba culture and transforms national cultural heritage into digital art design information [10], to realize the visual effect of Virtual character and the innovative development of Dongba culture [11].

2 Exploration of the Genes of Dongba Culture Based on the Culture Meme Theory

2.1 Development Status and Innovation Dilemma of Dongba Culture

The origin of the Dongba culture is closely related to the origins of several southwestern ethnic minorities. It is most closely related to the Dongba religion of the Naxi people. The Dongba religion and the Dongba scriptures were born out of the needs of folk life at that time. Therefore, Dongba culture also originated from the simple folk life, from all aspects of the daily life of the Naxi people. In the past, Dongba culture had long been passed down through Dongba scriptures, Dongba paintings, Dongba writings and oral narrations by Dongba priests. By the middle of the 20th century, the transmission of Dongba culture was seriously hindered. The literati and scholars set up a Dongba scripture translation group in the Lijiang County Cultural Center to carry out long-term translation and protection work on the preserved Dongba scriptures. With the joint efforts of the scholars, the preservation of Dongba culture has achieved significant results and laid a solid foundation for the study of Dongba culture.

With the development of Lijiang's tourism industry, Dongba culture has become one of the important pillars for the development of tourism industry in Lijiang. Nowadays, Dongba culture mainly relies on the tourism industry for dissemination and inheritance. Such a way of dissemination is a choice that conforms to the development of the region and the market, but it has limitations in time and space. Dongba culture relies too much on tourism industry and cultural and creative commodities, and its redesign content is very single, which is difficult to meet the needs of innovative development of Dongba culture. Excessive commercialization has also had a certain impact on the inheritance of the traditional and simple Dongba culture. In this paper, we combine the cultural modality theory and the hierarchical analysis method (AHP) to study the design transformation of Dongba cultural genes in Virtual character. It aims to systematically organize the typical

design factors in Dongba cultural genes, classify and evaluate the importance of these factors. Finally, it is applied to the Virtual character design to verify its rationality.

2.2 Dongba Cultural Genes Under the Culture Meme Theory

The basis of modal theory is derived from Darwin's theory of evolution. Its theoretical system is like the theory of evolution [12]. The difference is that modal theory is the "evolutionary theory" in the field of cultural evolution [13]. 1976, Richard Dawkins, a British biologist, proposed the term "Mimeme," which is shortened to "Meme," to correspond to the word "gene" in biological genes [14]. "Mimeme" abbreviated as 'Meme'. At the end of the 20th century. Chinese scholars also put forward the concept of "cultural gene", which is similar to that of biological gene [15]. Unlike genes in biological genetics, genes in cultural evolution cannot be directly subjected to scientific experiments [16]. Researchers mainly study it through historical and logical analysis. According to Liu Changlin, the underlying structures and ways of thinking that have had a profound impact on national culture and history are cultural genes [13]. In the field of cultural evolution, scholars generally agree that modality has the characteristics of direct copying, indirect imitation, carrying information, infecting memory and imitating evolution [15]. Therefore, the modality theory can play an important role in the process of extracting the design information of Dongba culture [17]. At present, there is no research content that uses modality theory to analyze the information of Dongba culture. In this paper, we hope that through the combination of the two, we can organize the gene map of Dongba culture applicable to Virtual character design.

3 Construction of an Information Gene Map of the Dongba Culture

3.1 Collecting Visual Information on Dongba Culture

We collect information on the visual patterns of Dongba culture. Firstly, we go to the birthplace of Dongba culture to conduct field research, communicate and discuss with local people and professionals who study Dongba scriptures, and take photographs of local Dongba texts and patterns. Secondly, combing through the relevant literature on Dongba culture, translating reports and research materials, and analyzing the composition system of Dongba culture information.

After the above investigation and analysis, we believe that the Dongba culture is roughly divided into three main parts: the fear and dependence on nature, which gave rise to nature worship. The need for and attachment to the family, resulting in ancestor worship. Imagination of the birth and death of life, resulting in the worship of multiple deities. Most of the visual information in Dongba culture is generated around these three components. Dongba culture genes are categorized as explicit and implicit. Generally, information that can be directly obtained through visual observation belongs to the dominant factors, such as shapes, patterns and colors. On the other hand, the elements of Dongba culture such as ritual traditions, living customs and spiritual connotations belong to the recessive factors, which are information that can only be obtained through a deeper understanding of the living customs and spiritual concepts of the Naxi people.

Through field research and literature analysis of Dongba culture, this study selected the core cultural genes that accompanied the development of Dongba culture and categorized them into natural pattern factor, character appearance factor, color factor, and semantic factor. We visually present the information categorization of these types of factors in the form of hierarchical structure mapping [18]. Finally, the cultural factors were transformed into design information that could be used in the Virtual character through AHP hierarchical analysis. The hierarchical model of Dongba cultural factors is shown in Fig. 1.

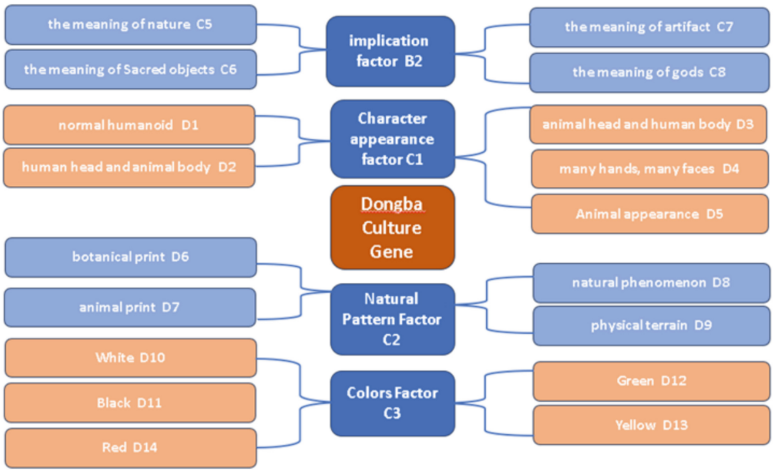


Fig. 1. Factor map of Dongba culture.

3.2 Semantic Factor Extraction

The visual message of the Dongba culture incorporates elements of Tibetan and Chinese cultures based on the indigenous animism. Its shape, pattern and color reflect the unique understanding and ideological perception of the ancient Naxi people on life, world and death. The meaning and information behind the visual images of Dongba culture are extracted based on the natural pattern factor and Dongba hieroglyphics, which were sorted out above, as typical representatives. They can be seen in Table 1.

3.3 Character Morphology Factor Extraction

The character form factor in Dongba mythology is the most typical dominant visual gene in Dongba culture. Besides, we can also extract the "genetic characteristics" of Dongba deities as usable design information. There are many types of Dongba deities, which are categorized in various ways, and the number of figures under each type is large. The deities are not closely related to each other, unlike the ancient Greek system of deities, which had a clear bloodline connection. Based on their appearance characteristics, we

have distilled the appearance characteristics of Dongba mythological figures into five types: normal humanoid, human-head-animal-body, animal-head-human-body, multi-handed and multi-faceted, and animal-form. These five character appearance factors run through the historical development of Dongba culture. No matter which stage, these

Table 1. Dongba Culture Genetic Information Meaning Factors.

Code	Prototypes Of Factors	Meaning of factor	Cultural Genetic Information
natural meaning C5		Elements arising from the concept of nature worship. Almost all the information about the Dongba culture can be found in the nature connotations in its roots.	The local animals, plants, mountains, rivers, lakes, and forests are all seen as having souls.
Sacred Meaning C6		In the development of the Dongba culture, things were gradually given sacred concepts. For example, this pillar of heaven, which is worshipped, is derived from the concept of "sacred mountain".	The sacred mountains, the sacred sea and the sacred trees are good and capable of bringing happiness and security.
Artifact Meaning C7		Sacred artifacts are a secondary concept to sacred objects. Most come from the weapons or symbols of the mythical Dongba deities. Sacred artifacts are believed to dispel evil and bad luck.	White conch, dharma knife, dharma bell, staff, etc.

(*continued*)

Table 1. (*continued*)

Deity Meaning C8		Numerous deities arose under the system of polytheistic worship. These deities differed greatly in form, character, and power.	In the eyes of the Naxi, there are good, neutral and evil Dongba deities. They respect or fear the spirits and want the good ones to protect them from the evil ones.

five appearance factors are present in the character traits of Dongba culture, and they are important visual information for the worship of multiple deities. We have extracted the factors of Dongba figure forms through field research and the book "Genealogy of Dongba Deities" (2000), which is shown in Table 2.

3.4 Natural Pattern Factor Extraction

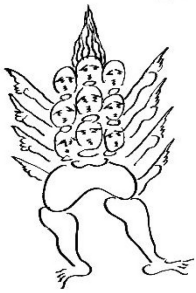
The Dongba culture is rich in symbols. Most of these symbols are derived from Dongba myths and are based on plants, animals and natural phenomena in the natural world. After the stories gave these natural objects certain meanings, they gradually evolved into decorative natural patterns, which became part of the genes of Dongba culture. Through

Table 2. Morphological factors of Dongba mythological characters.

Code	Prototypes of factors	Character Factor Pattern	Cultural Genetic Information
Humanoid D1			Toward the end of its development, the Dongba mythology was largely influenced by Indian and Tibetan cultures. It gave rise to many deities with the same appearance as human beings. This appearance is one of the "genes" of the

(*continued*)

Table 2. (*continued*)

human-head-animal-body D2			Dongba culture. In the early stages of the development of the Dongba mythology, many of these deities had distinct animal features. Most of these deities, represented by the deity of summer heat, are in the form of a combination of a human head and an animal body. Some of them also had the combined form of an animal head with a human upper body and a snake lower body. Such appearance is one of the "genes" of the Dongba culture.
animal-head-human-body D3			In the early stages of the development of the Dongba mythology, many of these deities had distinct animal features. Most of these deities were in the form of a combination of an animal head and a human body. This appearance is one of the "genes" of the Dongba culture.
multi-handed and multi-faceted D4			The fact that a deity has more than one head or more than one hand is also an important characteristic of Dongba deities. The fact that a character has multiple heads and hands is one of the "genes" of the Dongba culture.

(*continued*)

Table 2. (*continued*)

animal-form D5			Deities in animal form are also very common in the Dongba culture. These animals are either born as deities or become deities later in life. The animal form is also one of the "genes" of the Dongba characters.

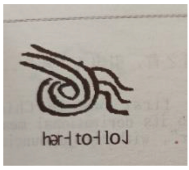
sorting, the nature pattern factors are categorized into four types: plants, animals, natural phenomena, and natural terrain. These four types of natural patterns originated from the natural environment in and around the Naxi settlements. They run through the development of the whole Dongba culture. The patterns with high frequency of occurrence in Dongba myths and Naxi costumes are selected for extraction, in Table 3.

Table 3. Natural pattern factors of Dongba culture.

Code	Prototypes of factors	Factor Pattern	Cultural Genetic Information
botanical pattern D6			Factor is derived from the wall paintings in the Dongba temples. Factor comes from the word "flower" in the Dongba hieroglyphics. It embodies the worship and reverence for nature.
animal pattern D7			The prototype of the factor comes from the wall painting of the 12 Chinese zodiac signs in the Dongba culture. Factor comes from the word "goat" in the Dongba hieroglyphics. Animals play an important role in the Dongba culture. Therefore, when designing Dongba characters, animal totems are also important cultural "genes".

(*continued*)

Table 3. (*continued*)

natural phenomenon D8			The factor archetype is derived from the natural element "wind". Factor is derived from the Dongba hieroglyphic word for "whirlwind". Natural phenomena such as wind, thunder and lightning are important visual messages of Dongba culture. They extend the corresponding ghosts and deities or motifs.
natural landscape D9			Factor archetypes are derived from the geographical features of the Naxi gathering places. The factor comes from Dongba hieroglyphics. Some natural landscapes were gradually influenced by nature worship and became visual information of Dongba culture.

3.5 Color Factor Extraction

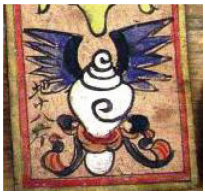
The color factor of Dongba culture is distilled from Dongba paintings. Born in the folk rituals of the Naxi people, Dongba paintings are paintings composed of lines and bright colors. During the rituals, the Dongba priests would draw the images of ghosts and spirits on bamboo or wooden boards, insert them into the soil, and arrange them into formations. The most prestigious large-scale scroll painting, "Shen Lu Tu", is 15 m long and filled with more than 300 images of deities and spirits, ghosts and phantoms, as well as nearly 80 images of sacred animals. The colors in Dongba paintings are layered and rich. In addition to the visible color information in the Dongba paintings, the color distinctions of many important things are also described in the Dongba myths, such as the color of the sacred mountains, the color of the sacred trees, and the color of the sacred stones. These colors are filtered and mainly categorized into five colors, which are white, black, green, yellow, and red, in Table 4.

4 Dongba Culture Genetic Information Evaluation System and Weighted Ranking

4.1 Dongba Culture Genetic Information Evaluation System

After extracting and classifying the genes of Dongba culture, to better provide design reference information for the Virtual character, we need to analyze the extracted Dongba culture genes. In this step, we adopted the Analytic Hierarchy Process (AHP) as the

Table 4. Color factors of Dongba culture

Code	Prototypes of factors	Factor color	Cultural Genetic Information
White D10			The prototype of the factor comes from the "white conch", a sacred object of the Dongba culture. In the early days of the Dongba culture, the color black was revered, and then slowly changed to the color white. White color represents light, daytime and goodness in the Dongba culture.
Black D11			The factor is based on one of the many "Shu" deities. Generally, the color black represents darkness, night, and viciousness in the Dongba culture. Generally dangerous Dongba figures will come with black or dark color. In the designer's eyes, black color can reflect the mystery of Dongba culture.
Green D12			The prototype of this factor comes from the mural of the Dongba War Deity. In the Dongba culture, the color green represents not only plants, but also snakes, bronze, and toxins. Some Dongba deities have a green skin color, partly reflecting their natural attributes and partly reflecting their fierceness.
Yellow D13			The factor is derived from the base color of the murals of the Dongba deities. Yellow color is mostly used as a symbol of deity, warmth and light in the Dongba culture. Most of the vestments of the Dongba priests are yellow.

(*continued*)

Table 4. (*continued*)

Red D14			The factor is derived from the murals of the Dongba deities. In the Dongba culture, red has bidirectional attributes like green. It represents both blood and danger as well as majesty.

measurement method to analyze cultural genes. AHP analysis was first proposed by American operations researcher T.L. Saaty in the 1970s [19]. It involves dividing the factors related to decision-making into several levels, subsequently making comparative judgments on the importance between two and two factors, establishing a judgment matrix, and calculating the weight of each element in the judgment matrix. In the process of constructing the factor map of Dongba culture, it is necessary to pay attention to the complexity and difference of cultural factors of the same type. To facilitate the extraction of the cultural information required for the Virtual character design, a hierarchical analysis method is used to compare the importance of each element with each other. We

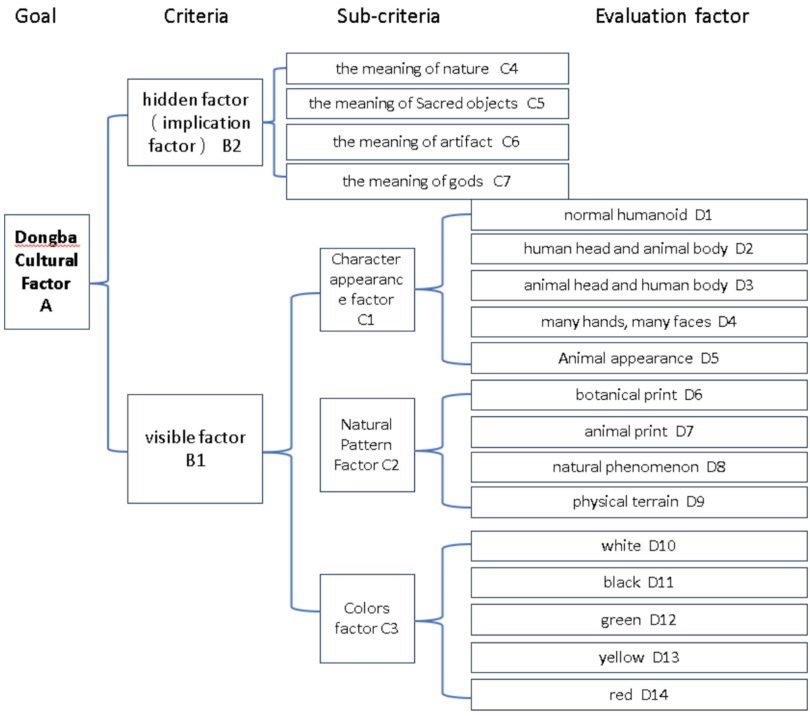


Fig. 2. Hierarchical evaluation model of Dongba culture factors.

calculate the relative importance of Dongba's cultural information in the design process as a basis for decision-making on the Virtual character design.

In the AHP model, the Dongba culture factor assessment system is divided into the target layer, criterion layer, sub-criterion layer and evaluation factor layer. The Dongba cultural factors correspond to the target layer and are coded as A. The explicit factors and implicit factors correspond to the criterion layer and are coded as B1 and B2, respectively; since there is only one type of semantic factor under the implicit factor B2, it is not possible to compare the scores of the two factors, so the four types of semantic factors, C4, C5, C6, and C7 are in the sub-criteria layer. The other three types of cultural factors correspond to the sub-criteria layer, coded as C1, C2, and C3, respectively, and each cultural factor and its information corresponds to the evaluation factor layer, coded as D1 to D18. Detailed information of the Dongba cultural factor assessment model is shown in Fig. 2.

4.2 Dongba Culture Factor Weighting Analysis

The Dongba cultural factor assessment model facilitates participants to rate different cultural factors of the same class based on the importance of the up and down factors. Based on the comprehensive analysis of the ratings given by multiple participants, we can filter out the most important cultural factors and design information for the Virtual character. Since the screened Dongba cultural factors were used in the domain of online Virtual character design, they were not used to study the importance of factors within the Dongba cultural domain. Therefore, in this evaluation, the trend of Virtual character and the main needs of Virtual character users were the reference for the participants' scoring.

Since in the field of regional and national culture, there are subjective differences in the perception of each element in the evaluation process. To minimize the errors caused by individual subjective factors, a multi-level evaluation system is constructed [19]. The evaluation system uses a scale of 1–9 for the questionnaire of the Dongba Culture Factor Weighting Questionnaire. Among them, the value 1 represents that one factor is equally important as another, 3 means slightly important, 7 means obviously important, and 9 means extremely important. Ten design professionals, five intended users of the Virtual character, and five design teachers were selected to form an evaluation team to score the importance of the Dongba cultural factors. The judgment matrix and evaluation weights of Dongba culture factor B layer is shown in Table 5, and the judgment matrix and weights of Dongba culture factor C layer are shown in Tables 6 and 7, and the judgment matrix and weights of each element in Dongba culture factor D layer are shown in Tables 8, 9 and 10.

When participants compare the factors two by two, there may be inconsistencies in the ratings. Therefore, it is necessary to conduct a consistency test on the results of the comprehensive assessment. When the consistency ratio $CR < 0.1$, it can be proved that the degree of data inconsistency is within the permissible range and passes the consistency test. When the consistency ratio $CR > 0.1$, it proves that the data do not pass the consistency test and need to be corrected. After testing each judgment matrix of the criterion layer, sub-criterion layer, and evaluation factor layer respectively, it is proved that the results of this evaluation have passed the consistency test and can be used as the

Table 5. Criterion level judgment matrix and weights.

	B1	B2
B1	1.000 0	4.000 0
B2	0.250 0	1.000 0
weights	0.800 0	0.200 0

Table 6. B1 judgment matrix and weights

	C1	C2	C3
C1	1.000 0	3.000 0	5.482 0
C2	0.333 3	1.000 0	2.287 0
C3	0.182 0	0.437 0	1.000 0
weights	0.653 7	0.235 4	0.110 9

Table 7. B2 judgment matrix and weights

	C4	C5	C6	C7
C4	1.000 0	0.333 3	0.607 0	0.361 0
C5	3.000 0	1.000 0	0.600 0	0.387 0
C6	1.647 0	1.667 0	1.000 0	0.646 0
C7	2.766 0	2.587 0	1.547 0	1.000 0
weights	0.122 8	0.214 2	0.257 6	0.405 4

Table 8. C1 judgment matrix and weights

	D1	D2	D3	D4	D5
D1	1.000 0	0.308 0	0.250 0	3.000 0	4.000 0
D2	3.251 0	1.000 0	3.003 0	8.032 0	4.664 0
D3	4.120 0	0.333 3	1.000 0	5.000 0	5.989 0
D4	0.333 3	0.124 0	0.200 0	1.000 0	0.600 0
D5	0.250 0	0.214 0	0.167 0	1.667 0	1.000 0
weights	0.145 8	0.447 9	0.294 7	0.046 7	0.064 7

basis for Virtual character design practice. The results of this consistency test are shown in Tables 11, 12, 13 and 14 respectively.

Table 9. C2 judgment matrix and weights

	D6	D7	D8	D9
D6	1.000 0	0.192 0	0.447 0	4.000 0
D7	5.213 0	1.000 0	5.999 0	6.998 0
D8	2.235 0	0.167 0	1.000 0	5.000 0
D9	0.250 0	0.143 0	0.200 0	1.000 0
weights	0.134 1	0.615 4	0.198 2	0.052 2

Table 10. C3 judgment matrix and weights

	D10	D11	D12	D13	D14
D10	1.000 0	0.615 0	0.448 0	1.261 0	0.404 0
D11	1.625 0	1.000 0	0.320 0	0.417 0	0.192 0
D12	2.232 0	3.125 0	1.000 0	1.747 0	0.293 0
D13	0.793 0	2.397 0	0.572 0	1.000 0	0.229 1
D14	2.473 0	5.216 0	3.412 0	1.000 0	1.000 0
weights	0.117 0	0.949 2	0.207 7	0.122 6	0.457 7

Table 11. Results of the criterion layer consistency test

Level	CI	RI	CR	Test results
Criteria	0.000 0	0.000 0	Null	Pass

Table 12. Results of B1 matrix consistency test

Level	CI	RI	CR	Test results
B1	0.003 0	0.520 0	0.005 0	Pass

Table 13. B2 Matrix consistency test results

Level	CI	RI	CR	Test results
B2	0.050 0	0.890 0	0.056 0	Pass

After the obtained data have passed the consistency test to prove their validity, the above sets of data are ranked. This leads to the ranking of the importance of Dongba

Table 14. Evaluation factor layer consistency test results

Level	CI	RI	CR	Test results
C1	0.083 0	1.120 0	0.074 0	Pass
C2	0.089 0	0.890 0	0.081 6	Pass
C3	0.080 0	1.120 0	0.071 0	Pass

cultural factors based on the needs of Virtual character design. The results of the weight ranking of Dongba cultural factors in Virtual character design are shown in Fig. 3.

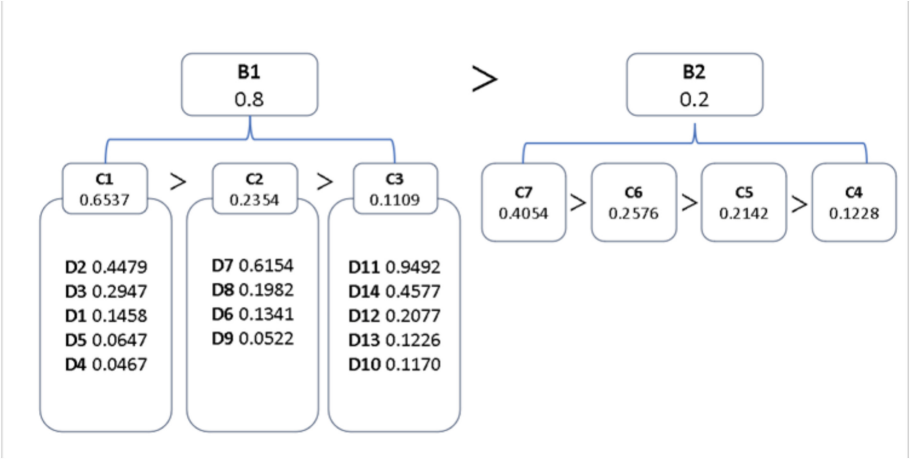


Fig. 3. Ranking chart of Dongba culture factor weights.

5 Design Practices of Dongba Culture in the Field of Virtual Character

According to the results of AHP hierarchical evaluation, we conclude that explicit factor B1 is much more important than implicit factor B2. Therefore, in the Virtual character design, the character form factor C1, the natural pattern factor C2, and the color factor C3 in the Dongba cultural genes are the main information considered. Among them, the weight of human head and animal body D3 in character form C1 is the largest. Therefore, in the shape part of the Virtual character, the fusion of human form and non-human form can be used. In the natural pattern C2, the animal pattern factor D7 is the most important information. This indicates that character designers and character users prefer animal patterns. Animal patterns can be used as visualization information of character character traits, presented in clothing, props or character modeling. In the color factor C3, black D11 is the color with the highest importance. This shows that

in the participants' conception, Dongba culture is characterized by mystery, secrecy, and darkness. More likely, this is because the current audience for Dongba culture is small and has primitive religious overtones. It is important to note that the five color factors were refined based on important elements of Dongba culture. These five color factors can represent the color characteristics of Dongba culture. Based on the contents of Dongba myths and the colors of the costumes of Dongba deities, it can be concluded that Dongba culture has undergone a transformation from revering black to revering white. The importance of white is higher in the system of Dongba religion, which symbolizes the role of goodness. However, in the whole Dongba culture gene, the remaining four-color factors have some importance. In the design of the Virtual character, based on the results of the comprehensive evaluation and the meanings given to the colors by the Dongba mythology, the main colors can be chosen according to the personality of the Virtual character designed [20]. In this practice, the weight evaluation ranking in C3 is used as the main color reference. In the implicit factor B2, the meaning of deities C7 is the most important cultural factor in B2. The hidden information behind the Dongba deities can be the priority reference content for Virtual character design, followed by the meaning of artifacts C6, the meaning of sacred objects C5, and the meaning of nature C4.

Based on the above analysis, we transformed the screened Dongba culture factors into visual information and put them into the process of Virtual character design. We successfully designed a Virtual character that meets the needs of users and at the same time conforms to the style of Dongba culture. The cultural information of this character is derived from the wind deity in the Dongba mythology and was designed with reference to the weights of the above cultural factors. Her specific content is shown in Fig. 4. This visual image meets the user's personalized needs. It also meets the Northeastern cultural and genetic characteristics in terms of content.

Fig. 4. Design Transformation of Dongba Cultural Factors.

6 Conclusion

The design of virtual characters has proved to be effective in transforming the information of Dongba culture based on the modal theory-AHP. The transformation of Dongba cultural genes into virtual images on the Internet allows regional ethnic cultures to spread across regions [21]. The combination of modal theory and AHP model can transform cultural genes into visual information more systematically and conveniently and establish systematic illustrations. These illustrations can help designers quickly translate cultural information into design works, especially when designers do not have in-depth knowledge of a particular culture. The methodology of this study has reference value and significance for the use of other cultures in virtual image design [22]. However, this study also has some shortcomings. Through the extraction of the genes of Dongba culture, we believe that it may be worthwhile to try to extract cultural genes with the help of rooting theory in this part of the process. This can provide a basis for the process of cultural gene extraction. As the characters in Dongba culture are complex and diverse, we would like to introduce artificial intelligence technology to organize the character genes and pattern genes of Dongba culture. This will help designers to obtain better and more detailed visual information.

References

1. Pyasik, M., Tieri, G., Pia, L.: Visual appearance of the virtual hand affects embodiment in the virtual hand illusion. Sci. Rep. **10**(1), 5412 (2020)
2. He, S., Song, Y.-Z., Xiang, T.: Style-based global appearance flow for virtual try-on. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022)
3. Park, S., Kim, S.P., Whang, M.: Individual's social perception of Virtual characters embodied with their habitual facial expressions and facial appearance. Sensors **21**(17), 5986 (2021)
4. Hepperle, D., et al.: Aspects of visual character appearance: self-representation, display type, and uncanny valley. Vis. Comput.Comput. **38**(4), 1227–1244 (2022)
5. Nelson, M.G., et al.: Walking in a crowd full of virtual characters: effects of virtual character appearance on human movement behavior. In: Advances in Visual Computing: 15th International Symposium, ISVC 2020, San Diego, CA, USA, 5–7 October 2020, Proceedings, Part I 15, Springer (2020)
6. Shen, D., et al.: Sound design of guqin culture: interactive art promotes the sustainable development of traditional culture. Sustainability **14**(4), 2356 (2022)
7. Mortara, M., et al.: Learning cultural heritage by serious games. J. Cult. Herit.Herit. **15**(3), 318–325 (2014)
8. Zhang, L., et al.: Ethnobotanical study of traditional edible plants used by the Naxi people during droughts. J. Ethnobiol. Ethnomed.Ethnobiol. Ethnomed. **12**, 1–16 (2016)
9. Qin, Z., Song, Y., Tian, Y.: The impact of product design with traditional cultural properties (TCPs) on consumer behavior through cultural perceptions: evidence from the young Chinese generation. Sustainability **11**(2), 426 (2019)
10. Nelson, H.G., Stolterman, E.: The Design way: Intentional Change in an Unpredictable World. MIT Press (2014)
11. Thompson, E.P.: Customs in Common: Studies in Traditional Popular Culture. New Press/ORIM (2015)
12. Tyler, T.: Memetics: Memes and the science of cultural evolution. Tim Tyler (2011)

13. Botz-Bornstein, T.: Genes, memes, and the Chinese concept of wen: Toward a nature/culture model of genetics. Philosophy East and West, pp. 167–186 (2010)
14. Blackmore, S.: Gene and meme. Encyclopedia of Animal Behaviour, Elsevier, London, pp. 67–74 (2019)
15. Chiao, J.Y., Blizinsky, K.D.: Culture–gene coevolution of individualism–collectivism and the serotonin transporter gene. Proc. Royal Soc. B: Biol. Sci. **277**(1681), 529–537 (2010)
16. Shifman, L., Memes in Digital Culture. MIT Press (2013)
17. Cardoso, G.C., Atwell, J.W.: Directional cultural change by modification and replacement of memes. Evolution **65**(1), 295–300 (2011)
18. Leigh, H.: Genes, memes, culture, and mental illness: Toward an integrative model. Springer Science & Business Media (2010)
19. Munier, N., Hontoria, E.: Uses and Limitations of the AHP Method. Springer (2021)
20. Ji, Z., Huang, W.-H., Lin, M.: Design mode innovation of local color cultures: a case study of the traditional female costume of Yi nationality. Designs **4**(4), 56 (2020)
21. Chen, S., et al.: My culture, my people, my hometown: Chinese ethnic minorities seeking cultural sustainability by video blogging. Proc. ACM Hum.-Comput. Interact. **7**(CSCW1), 1–30 (2023)
22. Castro, F.G., Barrera, M. Jr., Holleran Steiker, L.K.: Issues and challenges in the design of culturally adapted evidence-based interventions. Ann. Rev. Clin. Psychol. **6**(1), 213–239 (2010)

Research on the Cross-Cultural Promotion of Intangible Cultural Heritage "Xiangyunsha" Under Digital Empowerment

Xinnan Li and Yuanfang Zhao(✉)

College of Art and Design, Shenzhen University, Shenzhen 518060, People's Republic of China
zhaoyuanfang@szu.edu.cn

Abstract. Xiangyunsha, a dazzling treasure of traditional Chinese culture, carries a profound heritage spanning over a thousand years. As an intangible cultural heritage craft from the Lingnan region, Xiangyunsha has been exported overseas since the Ming Dynasty, praised by international friends as a "black shining pearl" of clothing art. It embodies the deep spirit and cultural essence of the nation and plays an increasingly important role in cross-cultural exchanges. How to fully explore the cultural implications embedded in Xiangyunsha and showcase its unique artistic charm is an urgent research topic. This effort not only serves as an important foundation for building cultural confidence but also as a symbolic cultural representation of the nation's identity. In China, intangible cultural heritage projects have garnered significant social attention, frequently appearing in various promotional activities. However, beyond the Lingnan region, the awareness of intangible cultural heritage among the general public, especially school-aged youths, remains low, with limited frequency and means of engagement. In-depth market research revealed that public awareness of Xiangyunsha is still vague, compounded by the vast, complex network information, including misleading content, making it difficult for people to understand the true nature of Xiangyunsha and differentiate authenticity.

Leveraging digital technology, this study aims to enhance the intercultural dialogue surrounding Xiangyunsha, fostering mutual learning and exchange between different cultures while promoting cultural confidence. With the rapid development of digital technology, a new stage has emerged, offering unprecedented opportunities for the promotion and evolution of traditional culture. In this context, utilizing the communication efficiency of digital empowerment for the cultural dissemination of Xiangyunsha represents not only a wise choice aligned with the times but also holds profound practical and historical significance, injecting new vitality into the inheritance and development of Chinese culture.

Through field investigations, experiments, and literature reviews, this paper conducts an in-depth study of Xiangyunsha's production techniques and market, revealing current issues. Building on this, the study analyzes the advantages of Xiangyunsha's digital development from the perspective of digital empowerment and explores innovative pathways for designing a digital platform for the intangible cultural heritage of Xiangyunsha. The research design includes the following aspects: (1) Establishing a database of related knowledge, colors, etc., to provide rich material and data support for the digital dissemination of Xiangyunsha; (2) Using existing digital platforms for educational promotion, summarizing and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 76–88, 2025.
https://doi.org/10.1007/978-3-031-94171-9_7

compiling existing issues to provide references for future improvements; (3) Conducting user research to create user profiles, determining the content direction and design style for the platform to ensure it meets user needs; (4) Developing the platform framework based on the above steps, ensuring its practicality and usability. The purpose of this research and platform design is to provide new ideas and methodological support for the digital promotion and dissemination of Xiangyunsha, facilitating its cross-cultural exchange and promoting its inheritance and development.

Keywords: Digitalization · Cross-Cultural Exchange · Xiangyunsha · Promotion

1 Introduction

Xiangyunsha dyeing technique is a traditional handmade craft originating from Shunde District, Foshan City, Guangdong, and is recognized as one of the nation's intangible cultural heritages. Also known as "Shai Lang" or silk dyeing technique, it is a long and intricate process in which a plant dye extracted from Shu Lang is used as the dye material and 100% silk serves as the sole substrate. Through repeated cycles of dyeing, sun-drying (referred to as "Shai Lang"), and a subsequent blackening process, the silk is transformed into Xiangyunsha [1]. Silk that has not undergone the blackening process but has been repeatedly dyed and sun-dried is termed Hongyunsha. When cotton or linen is used instead of 100% silk, the resulting fabric is not considered authentic Xiangyunsha; it is instead called Xiangyun cotton, Xiangyun linen, or simply fabric processed by the Xiangyunsha dyeing technique.

The history of Xiangyunsha can be traced back 1,700 years and is deeply rooted in the fertile cultural soil of Lingnan. It is a fabric unique to the Lingnan region of Guangdong since ancient times [2]. Since the Ming dynasty, Xiangyunsha dyeing—a representative intangible cultural heritage of Lingnan—has enjoyed international renown and is hailed as the "black shining pearl" in apparel art. It embodies the profound spiritual and cultural essence of the nation and its people. However, due to the complexity of the technique and its reliance on oral transmission, its inheritance is challenging. Furthermore, a lack of market guidance and promotion has led to a break in transmission, with varying degrees of understanding among students and facing dual challenges of heritage preservation and cultural recognition.

Based on the difficulties encountered by the Xiangyunsha dyeing technique, in-depth research on the process, and an analysis of the content on related intangible cultural heritage promotional platforms, this paper constructs a framework for a Xiangyunsha cultural promotion platform through typical case studies and student user surveys. The aim is to promote and disseminate the cultural value of Xiangyunsha among student groups. The research covers the following aspects:

1. Analysis and summary of the current content and issues in intangible cultural heritage platforms.
2. Learning and experimental practice of the Xiangyunsha dyeing technique to establish a color library and a knowledge base of the craft.

3. Development of a schematic diagram for the platform content construction process based on student demand surveys, thereby attracting student attention and promoting Xiangyunsha.

2 Research on the Cross-Cultural Promotion of Xiangyunsha Under Digital Empowerment

2.1 Current Status of the Dissemination of Intangible Heritage "Xiangyunsha"

Research indicates that the online promotion of Xiangyunsha primarily occurs on short video platforms. The quality of brand promotion varies considerably, and the authenticity of the information is often difficult to verify due to the lack of authoritative official promotion. Although documentaries exist, their limited reach has resulted in an age gap among those who understand the Xiangyunsha dyeing technique. Offline physical stores are mainly concentrated in Shunde and Hangzhou, with few in northern regions.

Visits to the Xiangyunsha museum and related exhibitions reveal that the presentation and transmission of Xiangyunsha culture are overly superficial, preventing visitors from perceiving its deeper value. This phenomenon is manifested in two main aspects: first, the exhibition content primarily focuses on showcasing the history and cultural development of Xiangyunsha; second, there are few explanatory and experiential activities, which fail to effectively reveal the true value of Xiangyunsha. Consequently, the public's understanding of the technique remains shallow, lacking deep reflection and emotional resonance. Establishing a digitally empowered promotion platform for Xiangyunsha is a critical step in transforming the technique from a static display into a dynamic digital entity. Through digital technology, the Xiangyunsha dyeing process can be permanently preserved online and reintegrated into everyday life, thereby achieving global dissemination and sustainable transmission [3].

2.2 Opportunities Brought by Digital Empowerment to the Xiangyunsha Cultural Platform

Salvage Protection: From "Memory Preservation" to "Dynamic Reproduction". Xiangyunsha faces serious inheritance challenges; its representative inheritors are generally over 50 years old, and the younger generation lacks both awareness and interest, leading to a severe shortage of skilled practitioners [4]. The complexity of the dyeing technique, which relies heavily on the experience and oral transmission of master artisans, makes it difficult to attract young talent. The advent of digital technology offers the possibility to archive and transmit the Xiangyunsha dyeing process with high precision. By recording the sun-drying operation, establishing dynamic databases, and employing virtual simulation and digital twin systems to deconstruct and simulate the process via VR, the trial-and-error cost for apprentices can be significantly reduced.

Breaking Boundaries in Communication: From "Niche" to "Mass Participation". By capitalizing on the wide reach and rapid dissemination capabilities of short video platforms, active engagement and the creation of topics related to Xiangyunsha as intangible heritage should be promoted. An immersive experience can be reconstructed

through AR technology, transforming perception from a two-dimensional presentation to a three-dimensional, immersive experience that deepens user engagement.

Educational Innovation: From "Master-Apprentice Transmission" to "Universal Learning". A gamified learning section should be established, creating an online experiential mode for Xiangyunsha production through mini-games and knowledge quizzes. This would allow users to appreciate the profound essence of the dyeing technique within a gaming environment. For example, Tencent's "Dunhuang Digital Worshiper" initiative enables users to learn about mural art and participate in restoration activities through a mini-program, with participation exceeding 30 million people [5].

Ecological Protection: From "Isolated Inheritance" to "Systematic Maintenance". A comprehensive database for the Xiangyunsha dyeing technique should be established to record key parameters—such as dye liquid concentration, dyeing techniques, sun-drying methods, and weather conditions—and to explain the underlying operational principles and factors affecting these processes. The dyeing process involves multiple variables, including dye concentration, light exposure, and water quality. Digital methods can be employed to optimize environmental parameters and provide real-time monitoring of dye concentration, air humidity, and temperature, thereby offering a scientific basis for the preservation of this traditional craft and promoting its sustainable development.

2.3 Case Analysis of Digitally Empowered Intangible Heritage Promotion Platforms

Case Analysis of China Crafts Network. China Crafts Network is a comprehensive website established by the national government to promote traditional Chinese crafts and intangible cultural heritage. It offers core content including news, craft demonstrations, an artisan work repository, trading services, educational training, and community interaction. The platform aims to connect intangible heritage inheritors with enthusiasts by providing an intangible heritage database, artisan certification systems, and policy guidelines, thereby serving craft exhibitions, transactions, education, and dissemination with both authority and professionalism. Although there are shortcomings in digital communication and user community engagement, the platform remains committed to constructing an ecosystem for traditional crafts, primarily through image-based interactions (see Fig. 1).

Case Analysis of the Dunhuang Silk Scarf Mini-Program. The Dunhuang Silk Scarf mini-program is a digital application developed collaboratively by Tencent and the Dunhuang Research Institute. It provides creative design and user customization experiences and comprises an official mall, a DIY design section, and a digital exhibition hall. Users can customize scarf designs by selecting Dunhuang elements for patterns and color coordination. Once designed, the creations can be saved, shared, and previewed for wearing effects. Although the mini-program performs well in terms of user experience and commercial consumption, there is a need to further deepen the cultural content related to Dunhuang and enhance user community interaction. Incorporating AR technology could further improve the realism of the wearing experience (see Fig. 2).

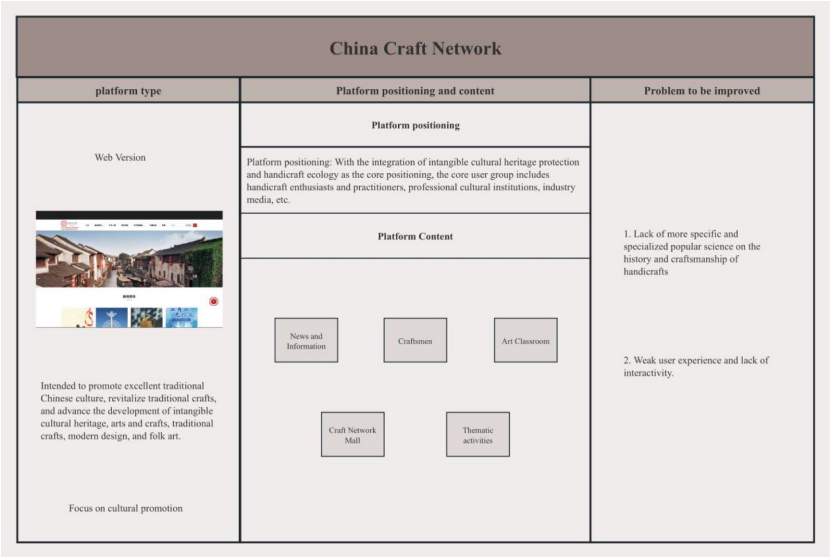


Fig. 1. Case analysis of China Handicraft Network.

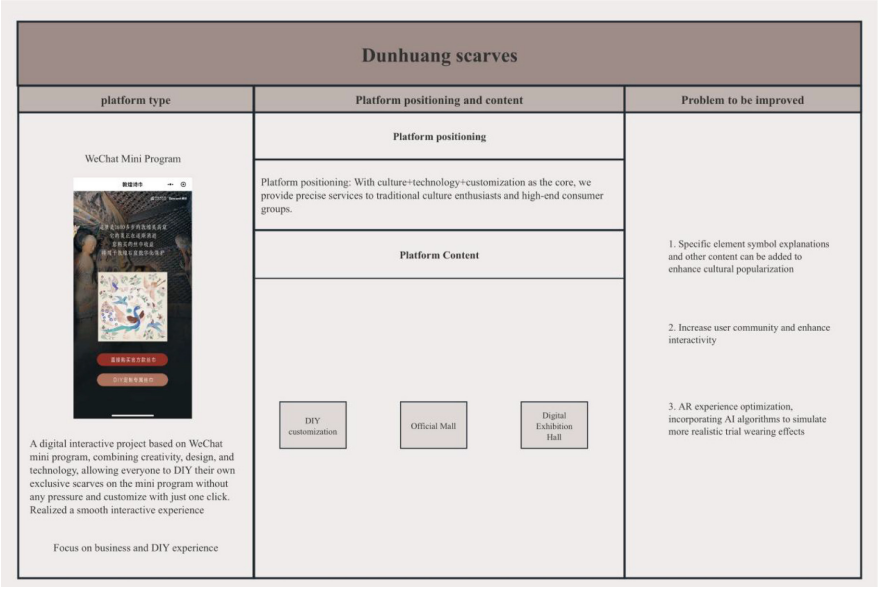


Fig. 2. Case analysis of Dunhuang Silk Scarf Mini Program.

Case Analysis of the Dongjia APP. Dongjia APP is a platform focused on artisans, offering an artisan ecosystem along with certification for intangible heritage inheritors and master studios. It features a cultural content community that showcases the

production of handcrafted items and the history of traditional techniques. The community has hosted events such as the "Intangible Heritage Redesign Competition," which attracted numerous users to engage in innovative embroidery pattern design; outstanding works have been mass-produced and received public acclaim. Dongjia APP places strong emphasis on the commercial transaction of artisan products, though its cultural content output is relatively limited. Currently, the platform primarily promotes craft culture through short videos and live streaming sales (see Fig. 3).

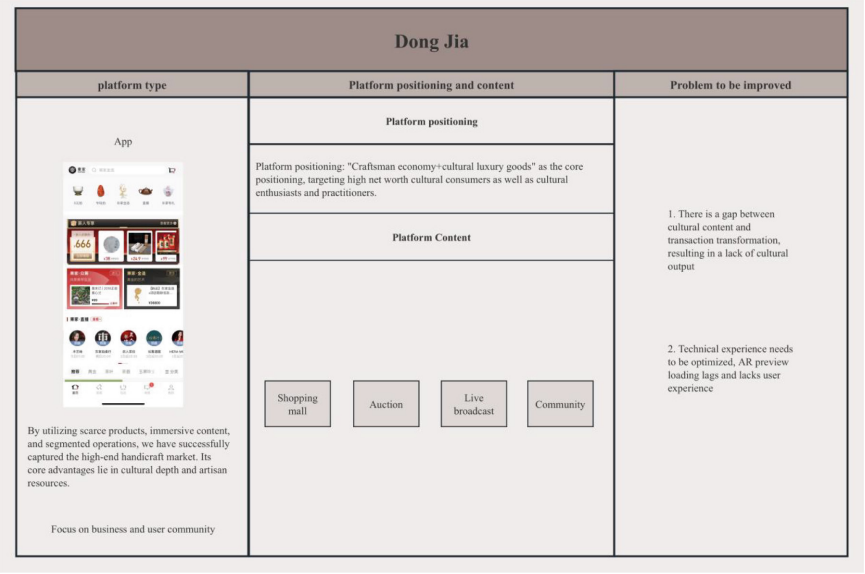


Fig. 3. Case analysis of Dongjia APP Figure.

3 Practical Exploration of Xiangyunsha Dyeing Technique

In-depth learning and practical experience are essential to fully comprehend the profound essence of the craft, which in turn supports the development of a color library and provides a basis for Xiangyunsha cultural knowledge Q&A.

3.1 Process of the Technique

The Xiangyunsha dyeing process is both intricate and meticulous. In addition, achieving a superior product requires the optimal combination of sunlight, temperature, and wind direction. The raw materials required include: raw silk fabric (high-quality 100% Bombyx mori silk), Shu Lang (used to extract the Shu Lang dye), river mud (sourced from rivers in the Pearl River Delta), and a flat grassy area (locally known as "padi rat" or "Hawaiian grass").

The specific process is as follows: preparation of raw silk, preparation of Shu Lang dye, immersion in the Shu Lang solution, sun-drying of the Shu Lang-treated silk, the blackening process, cleaning with river mud, fog spreading, cloth stretching, silk rolling, quality inspection and measurement, and finally, storage of the finished product. The operation is affected by weather conditions such as humidity, temperature, and wind direction; accordingly, the master artisan continually adjusts the concentration of the Shu Lang dye and the process steps as needed (see Fig. 4).

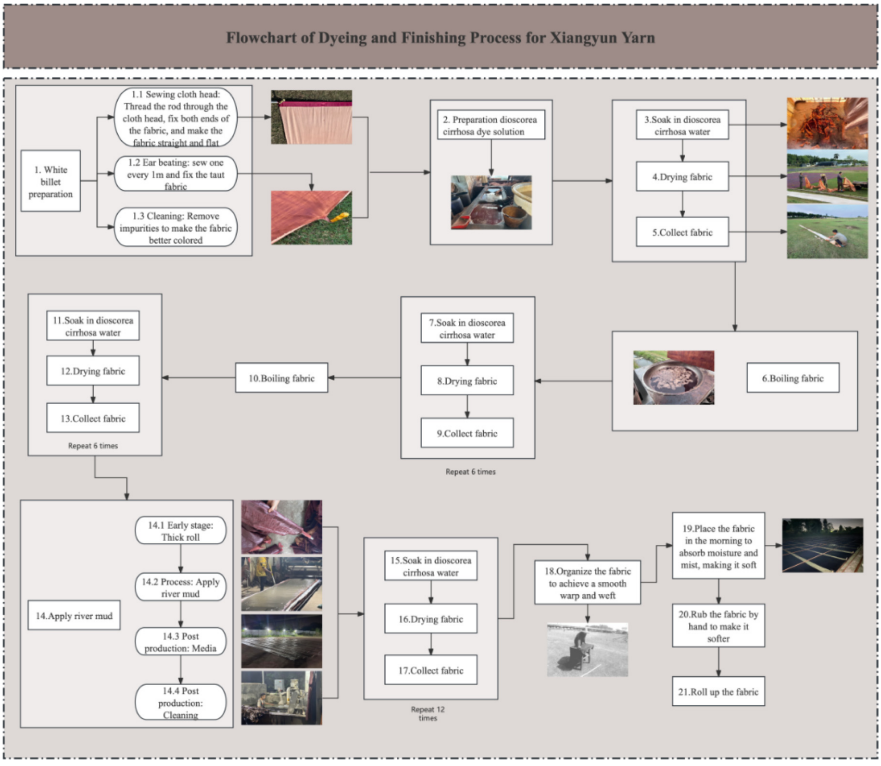


Fig. 4. Flow chart of Xiangyunsha dyeing and finishing techniques

3.2 Color Experiment

Building upon the traditional dyeing process, the present experiment also integrates plant dyeing with the sun-drying process. The experiment is divided into two parts: a basic experiment involving sun-drying and the blackening process, and an experiment combining sun-drying, the blackening process, and plant dyeing.

For the plant dyeing component, the amount of dye material is set at a ratio of 1:1 relative to the weight of the experimental fabric; the amount of mordant used is 0.15 times the fabric weight; the water used for dyeing is 20 times the fabric weight; and the

volume of the mordant solution is 0.3 times the fabric weight, diluted with water at a ratio of 30 times the fabric weight.

Throughout the Xiangyunsha dyeing process, factors such as temperature, sunlight, wind direction, and the concentration of the Shu Lang dye affect the number of repetitions of certain procedures. To reduce variables in the experiment, the dye concentration was standardized to that corresponding to the "Guo San Tou Shui" stage (see Fig. 5).

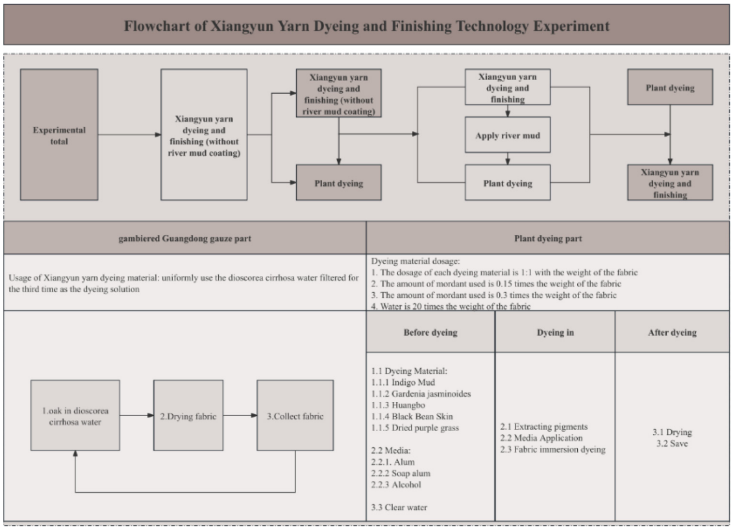


Fig. 5. Flow chart of Xiangyunsha color experiment.

The experiment successfully enriched the color palette of Xiangyunsha (see Fig. 6). Moreover, through "color digital empowerment" and "color visualization," the value of Xiangyunsha's colors was effectively conveyed. At various stages of the dyeing process, six color extractions were performed, with corresponding color numbers and codes assigned. This set of color libraries is applicable not only to the construction of a gamified color database but also holds significant value for physical printing, design creation, and process color reference, thereby establishing a model for digital empowerment of color [6].

4 Design Concept for a Digitally Empowered Cross-Cultural Promotion Platform for Xiangyunsha

4.1 User Analysis

This platform is targeted at university students. Currently, individuals with a high level of understanding of Xiangyunsha are generally over 35 years old and possess considerable economic strength; therefore, the focus shifts to the university student demographic.

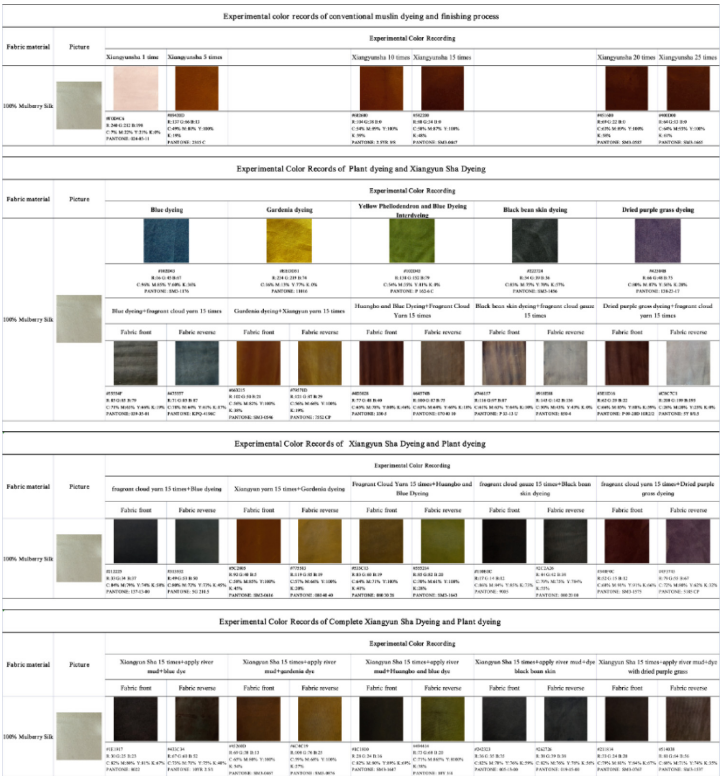


Fig. 6. Xiangyunsha dyeing and finishing experiment color library.

1. University students are likely to encounter topics on intangible cultural heritage through campus aesthetic and cultural courses or other social projects. Hence, the design of this platform will provide an exclusive digital empowerment platform for Xiangyunsha, characterized by high professionalism, strong community aggregation, immersive experience, and robust interactivity, thereby establishing an accurate cultural understanding of Xiangyunsha dyeing technique among students.
2. As a group characterized by vitality and creativity, university students can bridge the gap between intangible cultural heritage and modern society through "youthful expression," serving as an important link between tradition and the future. Their participation not only injects new vitality into intangible heritage but also plays an indispensable role in building cultural confidence.

User surveys indicate that the majority of students in South China have a basic understanding of Xiangyunsha, showing strong regional characteristics, while most students from other regions are largely unfamiliar with the technique. Over half of the student respondents are exposed to intangible heritage topics through their academic curriculum or projects, with their knowledge mainly derived from online sources, short videos, official documentaries, or professional experts. There are significant concerns regarding the professionalism and scientific accuracy of the information, and it is challenging

to contact professionals for interviews. There is a high demand for functionalities on the Xiangyunsha digital empowerment platform, including historical narration, technique comparison, expert contact, offline field visits, professional knowledge, and video tutorials (see Fig. 7).

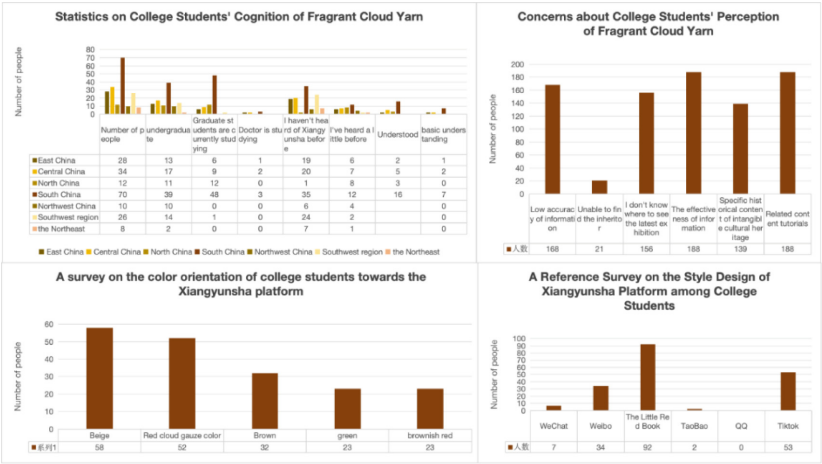


Fig. 7. Survey on college students' understanding of Xiangyunsha and Xiangyunsha platform design.

4.2 Design Concept and Showcase of the Digitally Empowered Xiangyunsha Cross-Cultural Promotion Platform

Platform Design Concept. Intangible cultural heritage is characterized by unique regionality and dynamic change, which necessitates shared communication, exchange, and discussion among people to reveal its cultural core. Therefore, establishing a digitally empowered platform is crucial for promoting intangible crafts. Specifically, a dedicated digital empowerment platform for Xiangyunsha dyeing technique can provide a targeted and professional cultural hub for students and enthusiasts. Based on preliminary user surveys and research on other digitally empowered intangible heritage platforms, the platform design primarily comprises sections for the homepage, intangible heritage artisans, a user community, and commercial displays.

Homepage Composite Section. The homepage should comprehensively present the platform's key and core content to attract users. This includes promotional videos of the Xiangyunsha production process, information on recent exhibitions and events, and an AR experience portal. By combining static text with dynamic AR experiences, the homepage will effectively showcase the charm of the Xiangyunsha dyeing technique. Additionally, integrating 3D imaging technology to display exhibition items will allow a full-dimensional presentation of the technique's details, thereby bridging the distance with users.

Intangible Heritage Artisan Section. The platform should invite intangible heritage artisans and master studios to join, utilizing live streaming and short videos to introduce the origins, technical principles, and authentication methods of the Xiangyunsha dyeing process. Through the sharing of authentic stories and close interaction, the platform can strengthen the connection between artisans and users, conveying the spiritual value of intangible crafts. Furthermore, integrating mapping features to highlight authentic Xiangyunsha shops and museum locations will facilitate offline visits.

Community Section. A community section is essential for creating a virtual space that fosters user participation, interaction, and retention. A well-constructed community enhances platform engagement, enriches the ecosystem, and increases user loyalty. By establishing dedicated Xiangyunsha topics to stimulate sharing and by incorporating interactive mini-games or knowledge quizzes, the platform can improve users' knowledge of Xiangyunsha and reinforce their cultural identity.

Commercial and Exhibition Section. This section should integrate brands related to Xiangyunsha apparel and cultural products, as well as the Xiangyunsha digital empowerment color library. By leveraging commercial mechanisms, the platform can facilitate the customization of Xiangyunsha colors and showcase the economic, cultural, and collectible values of Xiangyunsha products. Detailed product pages should clearly indicate information on technique, material, dimensions, and cultural connotations, transforming cultural attributes into economic value, thereby enhancing visibility and accelerating horizontal dissemination throughout society (see Fig. 8).

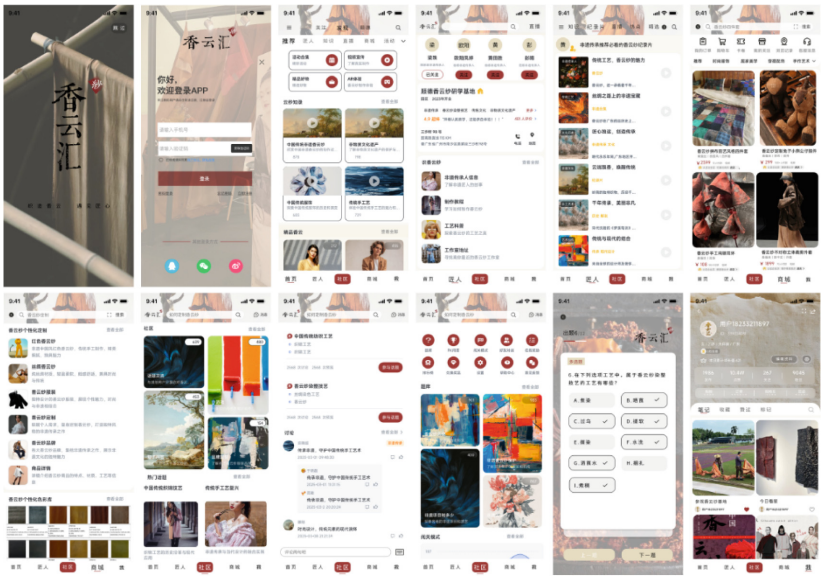


Fig. 8. Xiangyunsha digital empowerment platform UI design display.

5 Conclusion

Xiangyunsha dyeing technique is predominantly promoted through online short videos; however, the content is fragmented and monotonous, lacking scientific rigor. This results in a low level of understanding and participation among students. Based on the above research, it can be concluded that the digital empowerment promotion of the Xiangyunsha dyeing technique requires active collaboration from professionals and government agencies to establish a dedicated digital cultural platform. The incorporation of intangible heritage inheritors or specialized studios is necessary to ensure the platform's professional quality. Enhancing the digital visualization of Xiangyunsha's colors will facilitate users' comprehension of the technique's existing color profiles and support the commercial customization of these colors. This approach will improve the platform's educational, engaging, interactive, and disseminative qualities, thereby attracting a large student audience to actively learn about this intangible cultural heritage technique and addressing current challenges related to fragmented dissemination, limited understanding, insufficient experiential opportunities, and difficulties in authentication.

This paper actively explores the design of a digital empowerment platform for Xiangyunsha dyeing technique by analyzing case studies of other digitally empowered intangible heritage platforms. It provides insights into cross-cultural content development for Xiangyunsha, highlighting the importance of broadening its audience and protecting the craft. Future research should further investigate how digital platforms can better attract student groups, extensively promote Xiangyunsha dyeing technique, incorporate digitally empowered AR teaching technology, and enrich the digital color library data, thereby supporting the preservation and innovative transmission of this traditional art form.

Acknowledgements. The authors acknowledge the funding support by the Youth Fund of Humanities and Social Sciences Research of the Ministry of Education (Project No: 24YJCZH449), Guangdong Province Philosophy and Social Science Planning Project (Project Number: GD21YYS06).

References

1. IHChina Homepage. https://www.ihchina.cn/art/detail/id/14496.html. Accessed 25 Feb 2025
2. Li, W., Shi, Y.: Exploration of the adaptability of xiangyunsha fashion design. Decoration **07**, 127–129 (2008)
3. Su, H., Liu, M.: Design of a digital museum for intangible cultural heritage based on virtual reality technology: a case study of the Xiangyunsha digital museum. Ind. Des. **02**, 112–115 (2024)
4. Liao, Y., Wu, W., Pei, K., et al.: Exploration of promotion paths for intangible cultural heritage in the new media perspective: a case study of Xiangyunsha dyeing technique. Western Leather **46**(02), 52–54 (2024)
5. Zong, S., Liu, Z.: An analysis of the digital empowerment development path of traditional culture along the silk road: a case study of the dunhuang "digital worshiper" initiative. Publ. Perspect. **23**, 73–75 (2019)

6. Wen, X., Liu, C., Liao, X., et al.: Research on the visual digital preservation and inheritance of Xiangyunsha dyeing technique. Decoration **12**, 92–100 (2024)
7. Gao, Y., Li, Z., Yang, S., et al.: Research on the design and promotion strategy of a digitally empowered inheritance platform for "intangible craftsmanship" in Hebei. Commer. Exhib. Economy **21**, 129–132 (2023)
8. Li, A., Rao, J., Yang, F., et al.: Design and implementation of a digitally empowered display mini-program for miao intangible cultural heritAGE in Guizhou Province. Mod. Comput. **30**(21), 228–232 (2024)
9. Wu, B., Zhang, J., Cheng, N.: UI analysis and design of the intangible heritage impression APP. Mod. Comput. **30**(07), 101–104+108 (2024)

Research on Digital Interface Design for Celebrity Former Residence Museums Based on the KANO-AHP-TRIZ Hybrid Model: A Case Study of the Qi Baishi Museum Website

Na Liang, Bin Wang(✉), and YouLan Li

Hanyang University, Ansan 15588, Republic of Korea
wangbin19045@hanyang.ac.kr

Abstract. Thangka, also referred to as Tangka, With the continuous advancement of museum digital transformation, providing high-quality digital interactive experiences within limited resources and spatial constraints has emerged as a primary challenge for celebrity former residence museums. In response, this study proposes and validates a hybrid Kano-AHP-TRIZ model by integrating the Kano model, the Analytic Hierarchy Process (AHP), and TRIZ innovation principles. Focusing on the Qi Baishi Former Residence Museum as a case study, the research first employs the Kano model to identify and categorize user requirements, followed by the application of AHP to prioritize these key demands. Subsequently, TRIZ is utilized to address the technical and functional conflicts in the design process, and the optimized solution is comprehensively evaluated through TOPSIS. The findings demonstrate that this hybrid model precisely identifies and effectively resolves core contradictions in digital interface design, leading to significant improvements in interaction efficiency, visual presentation, and cultural integration for museum websites. This not only enhances the user experience for younger audiences and art enthusiasts but also provides systematic methodological guidance for museums seeking digital upgrades. While the present study offers valuable insights into the digital development of the Qi Baishi Former Residence Museum and delivers a feasible paradigm for achieving user-centered interface innovation in other celebrity former residence museums, it is limited to a single case. Future work should extend the model's application scope to various contexts and further refine the TRIZ methodology, thereby continuously promoting greater depth and breadth in the digital transformation of celebrity former residence museums.

Keywords: Celebrity Memorial Museums · KANO-AHP-TRIZ Hybrid Model User Experience Digital Interface Design

1 Introduction

With the rapid advancement of digital technology, cultural heritage preservation and dissemination have gained new opportunities. Digital interface design is crucial for enhancing the user experience and cultural communication value of museums [1]. In

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 89–105, 2025.
https://doi.org/10.1007/978-3-031-94171-9_8

recent years, digital exhibition technologies such as Virtual Reality (VR) and Augmented Reality (AR) have been widely applied in museum settings, achieving remarkable success, particularly in enhancing user immersion and experience [2]. However, small-scale celebrity former residence museums face unique challenges in digital transformation due to spatial and resource constraints. Despite their limited scale and financial support, these museums play an irreplaceable role in cultural heritage preservation and educational dissemination.

A critical research question arises: how can these museums achieve digital transformation with limited resources and space while meeting user needs and improving the visitor experience? Existing studies mainly focus on digital technologies, exhibition methods, or user satisfaction, but they lack systematic analyses of user requirements in museum digital interface design [3]. As user expectations rise, accurately identifying these needs and applying scientific methods for design optimization are key to improving user experience. In this context, theoretical approaches such as user requirement classification and priority assessment offer more targeted optimization strategies for digital interface design in museums [4]. This study combines the Kano model, the Analytic Hierarchy Process (AHP), and TRIZ innovation theory to propose a Kano-AHP-TRIZ hybrid model for systematically optimizing digital interface design in celebrity former residence museums. The Kano model classifies user needs into different levels, helping designers identify functions that are crucial for enhancing user satisfaction [5]. AHP offers a structured framework for prioritizing design elements [6]. Meanwhile, TRIZ addresses design conflicts, enabling the creation of innovative interface solutions [7].

Although these methods have been widely applied across different fields, there remains a lack of systematic research that integrates all three approaches in the context of digital interface design for celebrity former residence museums. Therefore, this study takes the Qi Baishi Memorial Hall as a case study to explore how the Kano-AHP-TRIZ hybrid model can be employed to optimize digital interface design, ultimately enhancing user satisfaction and engagement. Based on this framework, the study proposes the following hypotheses: **H1:** The Kano-AHP-TRIZ hybrid model can effectively enhance the design quality of the Qi Baishi Museum's digital interface, better meeting user needs and preferences. **H2:** By integrating TRIZ with the Kano model and AHP, this approach can provide innovative solutions for the digital interface design of the Qi Baishi Museum, thereby improving the overall user experience. **H3:** The existing conflicts between user requirements and design elements in the digital interface design of the Qi Baishi Museum can be effectively identified and optimized using the Kano-AHP-TRIZ hybrid model. This study aims to develop a hybrid model that integrates Kano, AHP, and TRIZ, using the Qi Baishi Memorial Hall website as a case study to explore the relationship between user needs and interface design. The proposed model seeks to provide a systematic theoretical framework and practical guidance for the digital transformation of small-scale celebrity former residence museums, ultimately fostering cultural dissemination and enhancing user satisfaction.

2 Literature Review

2.1 KANO-AHP-TRIZ Methodology

In today's digital and diversified environment, accurately capturing user needs is crucial for optimizing interface content and functionality. In the museum sector, the Kano model has been widely used to analyze digital functional requirements, assisting museum managers in prioritizing "basic needs" while allocating more resources to "attractive features" to enhance user satisfaction [3]. Liu et al. and Kapuria et al. conducted studies on the interactive design of the Minnan-Taiwan Museum and art museum websites, using the Kano model to analyze the expectations of users from different age groups and optimize interface design, significantly improving user satisfaction [8]. However, while the Kano model effectively classifies and ranks user needs, it has certain limitations in balancing different demand attributes and quantitatively evaluating individual preference variations [4].

To balance demand attributes, the Analytic Hierarchy Process (AHP) complements the Kano model by integrating qualitative and quantitative analysis. The Kano model identifies key functional indicators, while AHP ranks them by importance to optimize design. Studies show this approach enhances user satisfaction and engagement [9]. Wu et al. used the Kano-AHP model for smart jewelry design, improving user-friendliness and validating its effectiveness. However, evolving user demands and innovation challenges create conflicts between technical feasibility and functional needs, complicating design optimization.

TRIZ provides a systematic method for resolving technical contradictions in design [10] and has been widely applied across various fields [11]. Cheng et al. combined the Kano model with TRIZ to design a lower-limb exoskeleton, while Zhang et al. integrated AHP, QFD, and TRIZ to optimize combat robots, improving performance and user satisfaction. In museum digital interface design, TRIZ helps address conflicts like "information density vs. readability" and "technology vs. aesthetics" [1]. For celebrity former residence museums, it balances "cultural representation" and "functional usability," minimizing conflicts between function and artistic expression.

Research shows that combining the Kano model and AHP allows systematic classification and quantitative evaluation of user needs, while TRIZ enhances innovation, improving design efficiency and user satisfaction [3]. However, few studies have applied the Kano-AHP-TRIZ model to celebrity museum digital interfaces. Compared to traditional methods, this hybrid model offers a more precise and structured approach to meeting user needs and optimizing design. This study builds on previous research to develop and validate the Kano-AHP-TRIZ model, aiming to improve design quality and user experience in celebrity former residence museums.

2.2 Digital Interface Design for Museums

Digital interface design plays a key role in cultural heritage preservation and exhibition, enhancing visitor experience and cultural dissemination [1]. With improved internet infrastructure and mobile device adoption, museums use web-based platforms for online browsing, virtual interaction, and multimedia guides, overcoming the limitations

of offline exhibitions [12]. Compared to VR/AR applications and physical systems, web interfaces offer better cross-platform compatibility, maintenance, and scalability. They support various content formats, multilingual translation, personalized recommendations, and social media sharing, strengthening digital communication [3]. For celebrity former residence museums, web interfaces effectively present historical documents, artistic works, and interactive experiences. Their flexibility allows adaptive layouts and interactive features, meeting diverse curatorial needs and improving cultural communication. However, the Qi Baishi Former Residence Museum's digital interface has limitations in user experience, information structure, and interaction design (see Fig. 1). The current web interface does not fully meet user needs, requiring further optimization in visual presentation and functionality to enhance engagement and satisfaction.

Fig. 1. Existing Qi Baishi Former Residence Museum Web Interface.

2.3 User Experience Theory

User Experience (UX) reflects users' perceptions and attitudes when interacting with a product, covering usability, ease of use, and emotional value [13]. In museum digital interfaces, a good UX ensures smooth performance, enhances comprehension of exhibit information, and deepens cultural immersion. Meeting basic needs boosts satisfaction, while fulfilling attractive needs, like interactive creativity and personalized recommendations, increases engagement and retention [14].

Optimizing UX requires systematically classifying user needs and applying scientific methods to improve digital exhibition and interaction. This study uses the Kano model to categorize needs, AHP to quantify importance, and TRIZ to resolve design conflicts, supporting UX enhancement. While previous studies have explored the Kano model, AHP, TRIZ, and museum digital interface design separately, few integrate these methods for celebrity former residence museum web interfaces. This research develops a targeted digital interface optimization framework by combining user need classification, priority evaluation, and innovation-driven methodologies.

3 Method

This study takes the Qi Baishi Former Residence Museum as a case study, focusing on the design and optimization of its web interface. The research covers user needs investigation, design element analysis, development of innovative solutions, and implementation pathways.

By employing user interviews and literature analysis, the study first utilizes the Kano model to identify user functional requirements for the digital interface. Subsequently, the AHP (Analytic Hierarchy Process) is applied to prioritize interface design elements and identify key design issues. Finally, based on TRIZ theory and its inventive principles, this study resolves functional conflicts and optimization contradictions encountered in the interface design process, ultimately constructing a comprehensive optimization plan for the Qi Baishi Museum website. The final design is then evaluated and validated. The experimental procedure is illustrated in Fig. 2.

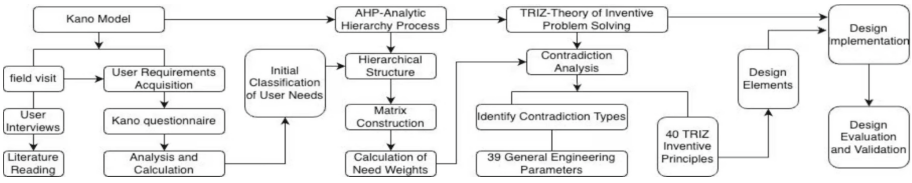


Fig. 2. Construction of the Digital Interface Design Process Based on KANO-AHP-TRIZ.

3.1 Constructing the KANO User Requirement Model

Kano Model. The Kano model is a user satisfaction-centered demand analysis method that classifies user needs through questionnaires or interviews. Based on the relationship between user satisfaction and requirement fulfillment, the Kano model categorizes user needs into five levels: Must-be Requirements (M): These refer to essential functions or design elements that users consider indispensable in a digital web interface. Their absence leads to strong dissatisfaction, but their presence does not necessarily enhance satisfaction. Performance Requirements (P): These represent expected functionalities that users wish to see in the Qi Baishi Former Residence Museum's web interface. The better they are implemented, the higher the user satisfaction. Attractive Requirements (A): These are innovative design features that exceed user expectations and provide a pleasant surprise. Their inclusion significantly enhances user satisfaction, but their absence does not lead to dissatisfaction. Indifferent Requirements (I): These are features that do not significantly influence user satisfaction, regardless of whether they are included in the interface. Reverse Requirements (R): These refer to design elements that users explicitly dislike or prefer to avoid. Their presence can lead to dissatisfaction, and they should be avoided in web interface design. By applying the Kano model, this study systematically identifies and prioritizes key user needs, ensuring that the digital interface of the Qi Baishi Former Residence Museum effectively meets essential expectations while incorporating delightful and innovative features.

User Needs Survey for the Qi Baishi Former Residence Museum Website. This study conducted a user needs survey on the digital web interface of the Qi Baishi Former Residence Museum, integrating the museum's practical characteristics and target user demographics. A combination of on-site investigations and user interviews was adopted for data collection.

The research team conducted an on-site survey and held discussions with museum management and technical staff to assess the existing web interface's functions and limitations. Additionally, 35 visitors with prior experience or interest in the museum's website were randomly interviewed in the exhibition area. Most users were aged 18–35 and focused on visual design, interactive experience, and functional improvements. The survey collected 78 user feedback entries, providing essential insights for requirement classification and interface optimization.

To refine requirement classification, this study reviewed relevant literature on museum digital interfaces and user experience. Experts from multiple fields evaluated and integrated the 78 collected user needs to ensure scientific validity. The panel included six interface designers, three visual communication professors, and three museum digitalization consultants. After expert review, 20 key user needs were identified. A two-level classification framework was established: Primary Classification: User needs were grouped into three dimensions—(A) Functional Needs, (B) Visual Needs, and (C) Emotional Needs. Secondary Classification: The 20 key needs were detailed into specific evaluation indicators (see Tables 1 and 2). This structured framework provides a foundation for optimizing the Qi Baishi Former Residence Museum's digital interface.

Table 1. User Demand Index System.

Experience Elements	User Needs
Function A	cheap price A1, easy to carry A2, rugged and durable A3, safety and environmental protection A4, education function A5, natural materials A6
Vision B	simple abstract B1, mysterious color B2, local specialty B3, high-resolution B4, creative and novel B5, color jump B6, fashionable and cool B7, traditional charm B8
Emotion C	fun interaction C1, relaxation and healing C2, meaning of praying C3, cultural connotation C4, lively and cute C5

Kano Attribute Classification and Requirement Acquisition. This study further applied the Kano questionnaire method to classify attributes based on user attitudes toward the presence or absence of specific interface features, as shown in Table 3. To ensure data representativeness and broad coverage, a combination of online and offline methods was used for the user requirement survey. A total of 135 questionnaires were distributed and 135 responses were collected, among which 126 were valid, yielding an effective response rate of 93.33%. For the online survey, 90 questionnaires were distributed via the "Questionnaire Star" platform, while an additional 45 questionnaires

Table 2. User Requirements Segmentation Table.

Number	User Requirement	Attribute	Number	User Requirement	Attribute
A1	Storage	Function	A11	Lumbar Support	Function
A2	Back Regulation	Function	B1	Warmth	Emotion
A3	Modularization	Function	B2	Security	Emotion
A4	Easier Installation	Function	B3	Interesting	Emotion
A5	Easy Washing	Function	C1	Organic Forms	Appearance
A6	Desk Board	Function	C2	Modern Style	Appearance
A7	Health Massage	Function	C3	Human-machine Engineering	Appearance
A8	Entertainment Function	Function	C4	Crude Wood	Appearance
A9	Easy Moving	Function	C5	Harmonious Proportion	Appearance
A10	Head Support	Function	C6	Composite Material	Appearance

were distributed through offline field surveys, ensuring the collection of valid responses. A total of 126 participants were surveyed, including 58 males and 68 females, primarily targeting the 18–35 age group. Among them, 78% were students and professionals in the cultural and arts sector, while 89% held a bachelor's degree or higher.

Table 3. KANO Questionnaire Design

Question		Scoring Criteria
Positive Question	When a certain attribute is realized, how do you feel?	1. Dislike
		2. Tolerate
		3. Neutral
		4. Accept as a matter of course
		5. Like
Negative Question	When a certain attribute is not realized, how do you feel?	1. Dislike
		2. Tolerate
		3. Neutral
		4. Accept as a matter of course
		5. Like

This study classifies user needs based on the Kano model and quantifies their impact on user experience using the Satisfaction Index (SI) and Dissatisfaction Index (DSI).

As shown in Fig. 3, the analysis results indicate that Expected Needs (O) play a critical role in enhancing user satisfaction. Specifically, fulfilling requirements such as easy cleaning (A5), warmth (B1), modern style (C2), and harmonious proportions (C4) can significantly improve user experience and should therefore be prioritized in the optimization process. Must-be Needs (M), including safety (B2), ease of installation (A4), and ergonomics (C3), represent users' basic expectations; failure to meet these needs may result in strong dissatisfaction. Indifferent Needs (I), such as fun elements (B3) and health massage functionality (A7), have a minimal impact on user satisfaction and can be considered as secondary optimization targets. Reverse Needs (R), such as B7, may reduce user satisfaction and should be avoided in the design process.

Number	M	A	O	I	R	Attribute	SI	DSI
A1	10	23	42	23	0	A	0.67	-0.34
A2	10	25	26	35	1	I	0.531	-0.365
A3	15	25	36	19	2	A	0.642	-0.421
A4	39	21	20	17	0	M	0.423	-0.619
A5	24	43	20	10	0	O	0.649	-0.69
A6	8	25	37	25	1	A	0.656	-0.343
A7	12	24	25	35	1	I	0.51	-0.375
A8	3	17	21	23	33	R	0.594	-0.313
A9	19	16	23	39	0	I	0.402	-0.361
A10	16	29	33	19	0	A	0.639	-0.464
A11	13	26	35	23	0	A	0.63	-0.402
B1	17	46	24	9	1	O	0.729	-0.656
B2	43	20	25	9	0	M	0.464	-0.649
B3	13	21	26	36	1	I	0.49	-0.354
C1	21	21	16	37	2	I	0.389	-0.442
C2	19	36	28	12	2	O	0.674	-0.579
C3	47	26	24	0	0	M	0.515	-0.753
C4	18	42	27	10	0	O	0.711	-0.619
C5	26	39	29	3	0	O	0.701	-0.67
C6	13	21	38	24	1	A	0.615	-0.354

Note: M. A, O, I, and R represent Must-be Requirement, Attractive Requirement, One-dimensional Requirement, Indifferent Requirement, and Reverse Requirement, respectively.

Fig. 3. User Requirement Attributes and Satisfaction Coefficients.

Overall, the findings of this study provide a scientific basis for product interface optimization and user experience enhancement. The optimization strategy should focus on key expected needs while ensuring that fundamental requirements are met to maximize user satisfaction. For the optimization of the Qi Baishi Former Residence Museum website, priority should be given to enhancing (A5), (B1), (C2), (C4), and (C5) to significantly improve user satisfaction. At the same time, it is essential to ensure that (A4), (B2), and (C3) meet basic user expectations to prevent dissatisfaction. In contrast, (A7), (B3), and (C1) have a less significant impact, and their optimization priority can be reduced. Meanwhile, B7 (Reverse Need) may negatively affect satisfaction and should be avoided in the design process. The optimization strategy should focus on key user needs, ensuring a rational allocation of resources to maximize user experience improvements.

3.2 Constructing the AHP Model

Construction of the Hierarchical Model for Celebrity Memorial Museum Website Requirements. To further clarify the core elements of the Qi Baishi Memorial Hall website interface design, this study introduces the Analytic Hierarchy Process (AHP) to

refine the results of the Kano model. As a systematic decision-making method, AHP is not only easy to use but also enables the structuring of hierarchical problems, calculation of criterion weights, and prioritization of alternative solutions, making it superior to other methods for requirement trade-off analysis [15].

Therefore, based on AHP, this study constructs a hierarchical structure model for the digital interface design of the Qi Baishi Museum. This model consists of: Goal Level: The optimized digital interface design elements (single objective). Criterion Level: Three dimensions of user needs—Must-be, Performance, and Attractive requirements. Sub-Criterion Level: 12 key requirements extracted from the Kano model analysis, covering Must-be, Performance, and Attractive needs. Additionally, these requirements are categorized into three primary dimensions: Functional, Visual, Emotional. This classification ensures scientific validity and rationality in the hierarchical model structure. The detailed framework is illustrated in Fig. 4.

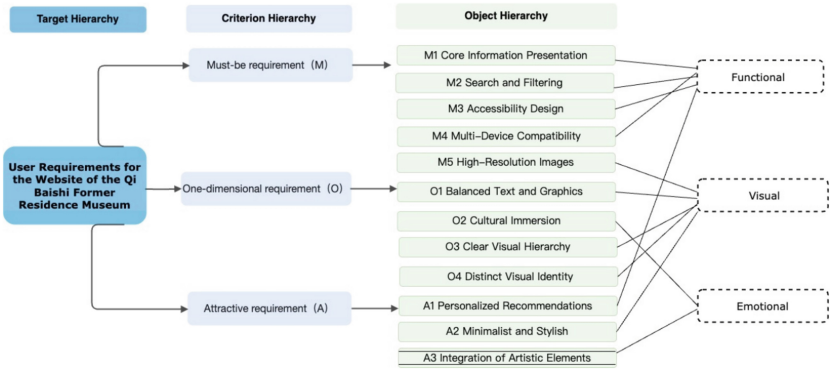


Fig. 4. AHP Model Construction.

Weight Analysis of Website Requirements for Celebrity Memorial Museums. To further identify the core elements of the Qi Baishi Memorial Hall website design, this study assembled an evaluation panel consisting of six experts, including three interface designers, two professors specializing in digital interaction design, and one museum user experience researcher. The expert panel applied Saaty's 1–9 scale method [16] to conduct pairwise comparisons of the requirement categories, thereby constructing a judgment matrix. The arithmetic mean method was then used to calculate the weight values of each requirement indicator. The final weight evaluation results are presented in Table 4. A consistency check was performed on the results, as shown in Eq. 1. If CR ≤ 0.1, the consistency test is considered valid.

$$CR = \frac{\lambda_{\max} - n}{(n-1) \cdot RI} \tag{1}$$

According to the calculations, the CR values for the M, O, and A indicators are 0.0789, 0.0907, and 0.0176, respectively, all of which are less than 0.1. Therefore, the consistency test is valid, confirming the effectiveness of the weight assignment in

the judgment matrix. Furthermore, the comprehensive weight values of the 12 sub-criteria were calculated, leading to the final ranking of the Qi Baishi Museum website requirements (Table 4).

Table 4. User Requirement Weights Table.

Number	1	2	3	4	5	6	7	8	9	10	11	12
Category	M	M	M	M	M	O	O	O	O	A	A	A
Requirement	M1	M2	M3	M4	M5	O2	O3	O1	A1	O4	A2	A3
Weight	0.3001	0.1587	0.0968	0.045	0.0327	0.0305	0.0828	0.1313	0.0159	0.0662	0.0254	0.0146

Among the three indicators in the criterion level, the Must-be attributes (M) have the highest weight proportion. At the sub-criterion level, the relative weight ranking of requirement elements, from highest to lowest, is as follows: M1 > M2 > M3 > M4 > O3 > O2 > M5 > O1 > A1 > B4 > A2 > A3. This comprehensive ranking provides key design guidelines for addressing critical issues related to the 12 requirement elements. However, during the design process, conflicts and contradictions between different design indices may arise. Therefore, it is necessary to apply TRIZ theory to resolve technical and physical contradictions effectively.

3.3 TRIZ Theory

Application of TRIZ Theory. Based on the Kano and AHP analysis, this study identifies 12 key requirements for the digital interface design of the Qi Baishi Memorial Hall. However, during the design implementation process, technical contradictions arise—optimizing one parameter may lead to the deterioration of another, ultimately affecting the overall design performance. To address these challenges, this study introduces TRIZ (Theory of Inventive Problem Solving) and applies its parameter contradiction analysis and inventive principles to optimize the interface design (Table 5). In the digital interface design of the Qi Baishi Memorial Hall, a systematic TRIZ-based analysis is conducted to identify key technical contradictions, determine the improving and deteriorating variables, and apply TRIZ's 40 inventive principles to develop targeted optimization solutions.

This study resolves key design conflicts using TRIZ-based strategies: Information Density vs. Visual Hierarchy: An information modularization strategy and key information amplification improve readability. Enlarging core text enhances clarity but may limit space. Hierarchical typography and contrast techniques balance information display, optimizing readability and retrieval efficiency. Functional Complexity vs. Interface Simplicity: A progressive information presentation strategy ensures an intuitive experience while maintaining functionality. Users access key details first, with more information revealed through interactions. Optimized typography and layout enhance cross-device compatibility, reducing visual overload and maintaining hierarchy clarity. Cultural Immersion vs. Interaction Efficiency: Color contrast and visual guidance techniques enhance artistic ambiance while maintaining clear information layers, preventing visual fatigue and improving user immersion. By systematically applying TRIZ,

Table 5. Application of TRIZ Theory.

NO.	Technical Contradiction	Improvement Variable	Worsening Variable	Related Invention Principles	Solution
1	The information density conflict between core content and visual hierarchy	23. Object Resolution	6. Fixed Object Area	1. Segmentation, 15. Dynamization	Adopt a modular display to highlight core information and optimize font size, spacing, and contrast for better readability
2	The interface complexity conflict between filtering and a minimalist modern style	17. Object Brightness	17. Object Brightness	3. Local Quality, 10. Preliminary Action	Optimize the filter box layout and interaction to help users quickly find content while keeping a minimalist interface
3	The readability conflict between accessible design and the integration of artistic elements	24. Object Complexity	24. Object Complexity	32. Color Change, 3. Local Quality	Use high-contrast colors and visual aids like icons and color blocks to improve readability and maintain aesthetics
4	The conflict in response speed between multi-device adaptation and personalized recommendations	23. Object Fineness	6. Fixed Object Area	15. Dynamization, 25. Adaptation	Use a progressive display to show personalized recommendations after users browse main content, preventing overload
5	The loading speed conflict between consistent visuals and high-resolution images	31. Object Manufacturing Adaptability	31. Object Manufacturing Adaptability	10. Preliminary Action, 32. Color Change	Optimize layout design to balance images and page composition, preventing high-resolution loading issues

(*continued*)

Table 5. (*continued*)

NO.	Technical Contradiction	Improvement Variable	Worsening Variable	Related Invention Principles	Solution
6	The conflict between cultural immersion and text-image balance in information hierarchy	31. Object Manufacturing Adaptability	31. Object Manufacturing Adaptability	3. Local Quality, 15. Dynamization	Use dynamic layering and fade effects to add depth to the cultural experience and prevent information overload
7	The detail representation conflict between high-resolution images and clear visual hierarchy	6. Object Size	6. Fixed Object Area	32. Color Change, 1. Segmentation	Adjust visual weight and add white space to highlight key images while keeping the interface clean
8	The conflict between text-image balance and artistic element integration	6. Object Size	31. Object Manufacturing Adaptability	3. Local Quality, 15. Dynamization	Use a grid layout to reduce interference and improve readability and hierarchy
9	The conflict between personalized recommendations and core information priority	23. Object Fineness	6. Fixed Object Area	25. Adaptation, 28. Mechanization	Use markers, color differences, and transitions to blend personalized content with core information
10	The user attention distribution conflict between clear visual hierarchy and cultural immersion	34.Internal Structure of the Object	31. Object Manufacturing Adaptability	31. Object Manufacturing Adaptability	Use gradient backgrounds and dynamic elements to guide attention and enhance cultural immersion while keeping information structured

(*continued*)

Table 5. (*continued*)

NO.	Technical Contradiction	Improvement Variable	Worsening Variable	Related Invention Principles	Solution
11	The aesthetic consistency conflict between a minimalist modern style and high-resolution images	23. Object Fineness	6. Fixed Object Area	3. Local Quality, 32. Color Change	Use a grid system and alignment to balance a modern minimalist style with high-resolution images
12	The adaptability issue between the integration of artistic elements and multi-device compatibility	12. Object Equivalence	24. Object Complexity	25. Adaptation, 10. Preliminary Action	Use responsive typography and adaptable visuals for optimal artistic display on all devices

this study enhances readability, balances visual and interactive elements, and improves user engagement, ultimately strengthening the museum's brand image.

Web Interface Design of the Qi Baishi Memorial Hall Based on TRIZ Theory. Based on TRIZ theory, this study first conducted a systematic design of the information structure for the Qi Baishi Memorial Hall website, as illustrated in Fig. 5. In the website navigation system, the primary navigation includes core modules such as Exhibitions, Visiting Information, Collections, Digital Experiences, and Interactive Education, ensuring that users can efficiently search for and access key information.

Fig. 5. Design Plan

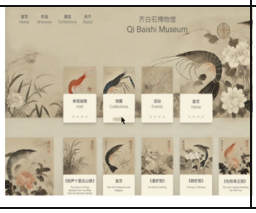
The collection display now features a refined classification structure, improving browsing and information retrieval. A new virtual exhibition hall offers an immersive cultural experience, enhancing user engagement and interactivity. This study optimizes the Qi Baishi Memorial Hall website's information architecture and functionality, meeting core user needs while improving interaction efficiency and usability. These enhancements enhance the user experience and strengthen the museum's digital communication capabilities.

4 TOPSIS Validation and Analysis

4.1 Understanding TOPSIS

TOPSIS is a ranking decision-making method based on proximity to the ideal solution. The fundamental principle of TOPSIS is to calculate the relative closeness of each evaluation object to the ideal solution, thereby distinguishing and ranking them in the decision-making process. The closer an evaluation object is to the optimal solution, the more it is considered the best option; conversely, the closer it is to the worst solution, the more it is regarded as the least favorable option. In this study, a structured evaluation system was employed to conduct TOPSIS analysis on three design schemes (A, B, and C) for the Qi Baishi Museum website, assessing and ranking them accordingly (Table 6).

Table 6. Evaluation Criteria and Ratings

Plan	A	B	C
Image			

4.2 TOPSIS Evaluation Process

Comprehensive Evaluation Index System for Design Schemes. To determine whether the final solution meets the needs of Qi Baishi Museum website visitors, a survey questionnaire was designed, incorporating research background and a rating scale. A 1–9 scale was used, allowing experts to independently evaluate the importance of each criterion. A total of 18 participants were invited to conduct the evaluation, including art designers, professors of digital media design, and experts in Qi Baishi's artistic heritage. During the evaluation process, the existing Qi Baishi Museum website interface (Scheme A) was compared with two newly proposed design schemes, Scheme B and Scheme C, as shown in Table 7. To ensure that the evaluation criteria accurately reflect the performance of each scheme, the evaluation indicators were established, and each scheme was rated accordingly. The comprehensive evaluation index for each scheme was then obtained, followed by an analysis of the consistency of expert opinions.

Normalized Decision Matrix. The Normalized Decision Matrix is typically used to address inconsistencies in the dimensionality of raw scoring data, ensuring that different evaluation criteria can be compared on the same scale (Table 8).

Using the vector normalization method, as shown in Eq. (2), each indicator is converted to a relative scale (typically ranging from 0 to 1 or normalized based on

Table 7. Evaluation Criteria and Ratings.

Design	M1	M2	M3	M4	M5	O1	O2	O3	O4	A1	A2	A3
A	7.2	6.8	7.5	6.5	6.0	7.0	6.9	6.8	7.1	6.4	6.2	6.5
B	7.8	7.5	8.0	7.2	6.8	7.3	7.5	7.6	7.9	7.0	6.8	7.1
C	8.5	8.0	8.8	8.0	7.5	8.2	8.3	8.1	8.6	7.8	7.5	8.5

Table 8. Normalized decision matrix.

Design	M1	M2	M3	M4	M5	O1	O2	O3	O4	A1	A2	A3
A	0.529	0.527	0.533	0.517	0.510	0.538	0.525	0.522	0.520	0.521	0.522	0.519
B	0.574	0.581	0.569	0.573	0.578	0.561	0.571	0.584	0.578	0.570	0.573	0.567
C	0.625	0.620	0.626	0.636	0.637	0.630	0.631	0.622	0.629	0.635	0.632	0.639

vector length). This ensures that all indicators maintain a consistent influence level in subsequent distance calculations, leading to more objective and rational evaluation results.

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}^2} \tag{2}$$

Determining the Positive Ideal Solution (A^+) and Negative Ideal Solution (A^-). After completing the normalization process, this study integrates the criterion weights obtained from the previous AHP analysis to construct a weighted normalized decision matrix. Finally, based on the evaluation attributes of museum website design, each criterion is classified as either a benefit-type or cost-type indicator. Using the weighted normalized values, the Positive Ideal Solution (A^+) and Negative Ideal Solution (A^-) are determined (Table 9).

Table 9. Ideal Solution.

Criterion	M1	M2	M3	M4	M5	O1	O2	O3	O4	A1	A2	A3
A + (Max)	0.188	0.098	0.061	0.029	0.021	0.083	0.019	0.052	0.042	0.010	0.016	0.009
A − (Min)	0.159	0.084	0.052	0.023	0.017	0.071	0.016	0.043	0.034	0.008	0.013	0.008

Calculating Relative Closeness and Ranking the Alternatives. After determining the distance between each alternative and the Positive Ideal Solution (A^+) and Negative Ideal Solution (A^-), the relative closeness (C_i) is calculated to assess the performance of each alternative. When C_i approaches 1, the alternative is closer to the Positive Ideal

Solution, indicating better overall performance. When C_i approaches 0, the alternative is closer to the Negative Ideal Solution, indicating poorer overall performance. Finally, all alternatives are ranked in descending order based on their C_i values, identifying the optimal and least favorable solutions (Table 10).

Table 10. Relative Closeness.

Design	Si+	Si−	$C_i = \frac{S_i^-}{S_i^+ + S_i^-}$
A	0. 195	0. 122	0. 385
B	0. 144	0. 152	0. 514
C	0. 103	0. 190	0. 649

From Table 10, Scheme C has the highest relative closeness value, making it the optimal choice. Scheme B ranks second, while Scheme A, the original design, scores the lowest.

Result Analysis and Discussion. TOPSIS results show that Scheme C outperforms the others in functionality, visual presentation, and cultural immersion, leading to higher user satisfaction and usability. Compared to Scheme A (the existing website) and the moderate improvements in Scheme B, Scheme C offers better multi-device compatibility, smoother interactions, and a more cohesive artistic style. Scheme B improves upon Scheme A in most areas but still has issues with loading efficiency and detail optimizations. While it integrates cultural elements better, it falls short of the ideal solution. If the museum seeks a cost-effective and quick upgrade, Scheme B could serve as a transitional option. Scheme A, the previous Qi Baishi Museum website, suffers from complex interactions, slow response times, and weak cultural representation, resulting in low user satisfaction and poor adaptability to new demands. TOPSIS analysis confirms that the Kano-AHP-TRIZ model-based Scheme C significantly enhances functionality, aesthetics, and interactivity. It effectively balances artistic and user experience needs, delivering a superior digital museum experience.

5 Conclusion

This study applies the TOPSIS method to evaluate digital interface designs for celebrity former residence museums, focusing on the Qi Baishi Museum website. Scheme C, developed using TRIZ theory, ranks highest, offering superior interaction efficiency, visual hierarchy, and cultural integration, meeting younger audiences' expectations. To support museum administrators, this study integrates the Kano model for user demand classification, AHP for design prioritization, and TRIZ for problem-solving, enhancing the reliability of the TOPSIS evaluation and providing a framework for digital transformation. We developed a user needs analysis model and a Kano-AHP-TRIZ optimization model to improve interactivity and innovation in museum websites. This approach resolves conflicts between user needs and functional design while enhancing cultural engagement.

Findings confirm that the Kano-AHP-TRIZ model effectively optimizes digital museum interfaces, improving user experience and providing a transferable framework for small-scale museums. However, since the study focuses on one museum, further validation across different museum types is needed. TRIZ applications in complex interface design also require refinement. Future research will expand datasets, test the model across more museums, and enhance TRIZ-based optimizations to ensure long-term applicability in cultural heritage institutions.

References

1. Li, J., Yu, J.: Preservation of intangible cultural heritage via digital technology: a systematic review. J. Cult. Herit. **52** (2021)
2. Tallon, L., Walker, K.: Museum 2.0: the transformation of museum exhibition and design. Curator: Museum J. **51**(1) (2008)
3. Xie, H., Chen, Z., Li, R.: Integrating KANO and AHP for user-centered museum exhibition design. Sustainability **14**(9) (2022)
4. Cao, Y., Luo, Z., Huang, W.: Application of AHP in cultural management decision-making: a case study of digital heritage projects. Inf. Manag. **58**(4) (2021)
5. Kano, N.: Attractive quality and must-be quality. J. Jpn. Soc. Qual. Control **14**(2) (1984)
6. Saaty, T.L.: Decision making with the analytic hierarchy process. Int. J. Serv. Sci. **1**(1) (2008)
7. Altshuller, G.: The Innovation Algorithm: TRIZ, Systematic Innovation and Technical Creativity." Technical Innovation Center (1997)
8. Kapuria, R., Gupta, N., Sharma, R.: Applying KANO model to assess user satisfaction in art museums: a multi-age group analysis. J. Arts Manag. **6**(2) (2020)
9. Zhao, R., Xu, L.: Modular furniture design for children's museums: integrating Kano, AHP and GRA. J. Intell. Manuf. **34**(1) (2023)
10. Ilevbare, I.M., Probert, D., Phaal, R.: A review of TRIZ, and its benefits and challenges in practice. Technovation **33**, 2–3 (2013)
11. Yao, K.C., Li, K.Y., Xu, J.R., Ho, W.S., Shen, Y.H.: Application of TRIZ in product innovation and design. Int. J. Innov. Sci. (2022)
12. Zhang, Y., Chen, L., Tang, Y.: Virtual reality integration in museum exhibitions for immersive experiences. Displays **59** (2019)
13. Hassenzahl, M., Tractinsky, N.: User experience - A research agenda. Behav. Inf. Technol. **25**(2) (2006)
14. Bresson, G., Mackenzie, R.: Using KANO model to evaluate tourist satisfaction and identify key factors for heritage site improvement. J. Cult. Herit. Manag. Sustain. Dev. **9**(3) (2019)
15. Liu, Y., Eckert, C.M., Earl, C.: A review of fuzzy AHP methods for decision-making with subjective judgments. Expert Syst. Appl. **161** (2020)
16. Yang, Y.P., Tian, H.L., Jiao, S.J.: Product design evaluation method using consensus measurement, network analysis, and AHP. Math. Probl. Eng. **2019**(1) (2019)

Design, Technology, and Applications of Gamified Exhibitions: A Review

Yanting Liao[1](✉) and Guo Jin[2]

[1] University of the Arts London, London SW1P 4JU, UK
liaoyant0127@163.com
[2] Sichuan Fine Arts Institute, Chongqing 28 401331, China

Abstract. In recent years, advancements in information technology and the growing need for cultural inquiry have gradually driven the emergence of gamification as a prominent trend in exhibition design. The novelty and innovation of gamification can attract new audiences to exhibitions. Its interactivity and reward mechanism can enhance the visitors' exhibition experience. Although some scholars have already conducted reviews on gamified cultural heritage, and museum gamification, the relevant content and details of gamified exhibitions also lack more systematic, in-depth, and comprehensive summaries. Therefore, this paper aims to provide an all-round review of gamified exhibitions from multiple dimensions, including application domains and scenarios, assessment methods and technologies, presentation formats, gamified content, assessment methods, and future development. Application domains and scenarios, e.g. cultural heritage, museums, education, archaeology. Gamification methods and technologies, e.g. Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR), Interaction Design, Mobile Applications/Devices, 3D Modelling, Interactive Narratives, Mini-Games/Complex Games, QR Codes/NFC, and Personalised Experiences. Presentation formats, e.g.: Virtual/Realistic exhibitions, temporary/permanent exhibitions, laboratory projects, additional/original games. Assessment methods, e.g.: questionnaires, interviews, observations, Short Flow State Scale, Think aloud method. This study clarifies the framework of gamification design in exhibitions and analyses in detail the elements included in gamification exhibition design from an overview perspective. The findings offer valuable references for the design and development of gamified exhibitions while highlighting the potential of gamification in shaping future exhibition practices.

Keywords: Gamification · Exhibition · Curation · Technology · Application domains

1 Introduction

As the "Soul of a museum experience" [1], exhibitions have evolved from a one-way knowledge transfer to a dynamic equilibrium between the curator, the artist and the audience. With this trend, gamified exhibitions, a new way of exhibition, are gradually emerging, offering potential solutions to address the convergence of multiple exhibition

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 106–116, 2025.
https://doi.org/10.1007/978-3-031-94171-9_9

purposes such as 'learning', 'education', and 'Immersive Experience' [2–5]. Gamification exhibition refers to the study of gamification in the context of exhibition, it aims to integrate "The use of game design elements and principles in non-game settings to improve user engagement" [6] into aspects of the exhibition's spatial design, narrative framing, and other curatorial strategies [3].It can enrich artworks and artists' creative ideas through innovative technology, enhance audience engagement, optimize the immersive efficiency of the exhibition space, and deepen the audience's cognitive memory [7–12].

Currently, the development of gamification covers multiple areas and presents a trend of multidimensional innovation [13]. This is particularly evident in the development and use of digital technology. For example, enhancing immersive visit experiences based on drone exploration game quests [14]. At the same time, gamified exhibitions encompass a variety of devices and technologies, such as serious games in which smartphones assist visitors in building personalized museums [3]; LIFEPLUS utilizes augmented reality (AR) technology for guided tours [15]. The Wieng Yong House museum creates a fabric-based virtual reality (VR) application [7]. 'The Ara As It Was' uses technology to integrate education and entertainment [12].

In the research of gamification exhibition, some scholars have conducted a review of related fields. Marques et al. [16] point out that the practice of gamification in the field of cultural heritage focuses on historical buildings and archaeological sites. Madsen and Maria [11] argue that gamification faces challenges such as a lack of a structured framework and conceptual confusion while stimulating motivation and shaping experience. Wang et al. [17] highlight the potential of digital preservation and the challenges of technology convergence by analyzing the role of gamification and immersive technologies in promoting the understanding of cultural heritage. Ćosović and Brkić [18] and Paliokas and Sylaiou [19] focus on the educational role of serious games, which can enhance learning, but there is a lack of framework and content disconnection. Falk et al. [20] explore the application of game technology in cultural heritage conservation and interactive experience.

Although there have been reviews of gamification related to museums and cultural heritage, most of them focus on specific contexts and lack systematic analysis and design discussion of exhibitions. Therefore, this paper conducts a systematic review on the theme of gamification exhibitions, aiming to provide a theoretical basis for the development of gamification exhibition. It covers multiple scenarios, different carriers, technology use, design framework, audience needs, implementation effects, and future trends. By synthesizing the existing research results on the application of gamification in different contexts such as museums, this study can provide a reference for the practical innovation and theoretical deepening of the application of gamification in exhibitions.

2 Application Domains and Scenarios

In recent years, gamified exhibitions have been widely used in museums for cultural heritage display, learning, education and archaeological exploration. They enrich the audience's experience in an interactive way and provide new possibilities for the dissemination of cultural heritage in the new era.

First, in terms of cultural heritage, as shown in Table 1, the specific applications f of gamification exhibitions in the direction of cultural heritage can be divided into three aspects: museum collections, cultural sites, and regional cultural history. Gamification offers new possibilities for the display of museum collections, allowing the audience to appreciate artworks from a multi-dimensional perspective and deepening their understanding of cultural relics and their cultural contexts. For example, the Museu Nacional Soares dos Reis [21]. For cultural sites, gamification not only enhances the audience's exploration experience and optimizes the modes of audience engagement, but also contributes to the protection of sites and ancient buildings. For example, The Arlaten Museum [14]. In addition, gamification also integrates local history, legends into the gamified experience, allowing the audience to deepen their understanding of a specific cultural context through interaction, such as Ulysses in Irish culture [22].

In the field of museum learning and education, gamification exhibitions promote the deepening of the educational function of museums at multiple levels. For example, educational scenarios [23], user experience [24], youth education [25], informal learning [26], and game-based learning [27]. In terms of archaeological exploration, gamification exhibitions cover archaeological culture [28] and archaeological steps [29], enhancing the audience's understanding and interest in archaeological work through interactive experiences.

Table 1. Application Domains and Scenairos.

Major categories	Subcategory	Museums	Example	Authors
Cultural Heritage	Museum collections	The Museu Nacional Soares dos Reis		[21]
Cultural Heritage	Ancient Agora	Ancient Agora of Athens		[30]
Cultural Heritage	Ancient remains	The Arlaten Museum		[14]
Cultural Heritage	Irish culture	Ulysses by Irish author James Joyce		[22]
Cultural Heritage	History of Portugal	History of Fado		[31]
Museum Learning and Education	User attraction	The world soil museum		[24]
Museum Learning and Education	Informal learning	Boat museum		[26]
Museum Learning and Education	Game-based learning	The Macedonian Museum of Contemporary Art		[27]
Museum Learning and Education	Local cultural knowledge	Boat museum		[32]
Museum Learning and Education	Museum Education	The National Palace Museum in Taipei		[23]
Museum Learning and Education	Teenage Experience	The natural history museum of Funchal		[25]
Archaeological	Archaeological culture	The Chimú culture from Peru		[28]
Archaeological	Archaeological steps	Argos Archaeological Museum		[29]

3 Gamification Methods and Technologies

Gamification exhibitions use gamification elements such as role-playing, decryption exploration, and immersive plot interaction, and combines badge collection, leaderboards, and other mechanisms to enhance the interactive experience. Technologically, through both emerging and traditional methods, cultural scenes and gamified stories are integrated to strengthen the audience's perception of cultural context and improve their interest in museum learning.

Gamification elements are one of the commonly used methods in gamified exhibitions. As shown in Table 2, the gamification element methods used in gamified exhibitions can be specifically categorized as role-playing games [33], collecting games [4], collection-unlocking games [34], scavenger hunts [35], charades [36], jigsaw puzzles [37], score leaderboards [8], and identification-type games [38]. The gamification elements are intended to draw on arrangements in games as a way to enhance the connection between visitors and the museum's collections and educational goals.

Table 2. Gamification Methods.

Major categories	Specific approaches	Results	Authors
Gamification elements	Crossword	Helping knowledge transfer	[36]
Gamification elements	Identification-type game	Enriching the visitor experience	[38]
Gamification elements	Collecting game	Providing new interactive ideas	[4]
Gamification elements	Scavenger hunts	Introducing excavations in Greece	[35]
Intelligent games	Image Categorization Game	Enhancing learning in museums	[43]
Digital games	Role-playing and Collecting game	Motivate students participation	[10]
Serious games	Role-playing and Collecting game	Promote youth learning	[41]
Serious games	Immersive experience	Increase Engagement	[40]
Gamification elements	Ranking	Motivates audience participation	[8]
Gamification elements	Jigsaw puzzles	Increase audience interest in learning	[37]
Serious games	Task challenges	Positively impacts learning	[42]
Gamification elements	Role-playing game	Reinforces willingness to learn	[33]
Mini-games versus Complex games	Monopoly and Quiz Games	Enhances audience knowledge	[44]
Serious games	Interactive narrative	Immersive learning	[39]
Gamification elements	Collection-unlocking game	Enhances user experience	[34]
Pervasive games	Puzzle	Enhances artistic understanding	[3]
Mixed reality games	virtual golf	Understanding of soil knowledge	[24]

Serious games are often used to capitalize on the characteristics of the museum or exhibition itself, specifically to complement and enhance the museum's role in education, cultural communication, and audience engagement [19]. For example, interactive narratives [39], immersive experiences [40], role-playing [41], task challenges [42].

In addition, other gamification approaches are covered in Table 2, such as intelligent games [43], mixed reality games [24], digital games [10], and mini-games versus complex games [44], pervasive games [3].

First, in terms of emerging technology methods, as shown in Table 3, the specific technologies used in the gamification exhibition can be divided into AR, VR, MR technology, such as, FabricVR [7], TeosGO [45], ARquiz [46], MuseumEye [47], The Ara As It Was [12]; 3D modeling, such as The 3D city game Ghent [48] project; NFC Technologies, such As MOMAP[49]; QR codes, such as gamified smart object platform [50]; applications, such as CrossCult [51], PAGANS game [52], DBpedia [53], Strike-a-pose[54], and AR Voyage [55].

It is worth noting that the technical strategies of gamification exhibitions are not limited to emerging technologies. Traditional gamification methods also show significant application value, for example, card media [56] and participatory design [5].

Table 3. Gamification Technologies.

Category	Technical name	Technical form	Software and hardware	Results	Authors
Application\QRcode	Platform	Smartphone	DBpedia\MongoDB\Python Flask\JavaScript\React Native		[50]
VR\Application	FabricVR	HMD	Unity3D\Blende\ RealityCapture		[7]
Application	CrossCult	Devices	Autotelic creation games		[51]
Application	PAGANS game	Web	Java \Tomcat \JavaScript		[52]
AR\Application\GPS	TeosGO	Mobile devices	visual-ization		[45]
QRcode\Application	DBpedia	HMD	PythonFlask\JavaScript\React native		[53]
MR\Immersive design	MuseumEye	Computer	Maya\ZBrush\Sense3D\Unity3D		[47]
VR\AR\Immersive design\3D Modeling	Ghent	Computer	Sketchup\Blender\Z-brush\JavaScript\Github		[48]
MR\Interaction design	The Ara As It Was		Gear VR viewer		[12]
Application	Strike-a-pose	Web	JavaScript\Python\TensorflowJS\MoveNet\Flask		[54]
AR\Application	AR Voyage	Mobile devices			[55]
Engagement design	Dress to Win	Documents			[5]
AR\Application/Interaction design	ARQuiz	Mobile devices	arce3\OpenMVG1\mongodb4\unity5		[46]
Application\Interaction design\NFC	MoMap	NFC	JetBrains\ \Java\MySQL		[49]
Interaction design	Touch Tomb		3D Print\LED		[56]

4 Presentation Formats

The gamification exhibition can be divided into permanent/temporary exhibition, virtual/physical exhibition, additional/original games and laboratory projects. Among them, additional and original games, as a gamification form of gamification exhibition, enhance the audience experience through multi-level interactive design.

First, in the permanent and temporary exhibition formats, as shown in Table 4, permanent exhibitions usually include museum collections, cultural sites, and individual gamified exhibition of certain artworks. For example, PP-IPTEK Museum TMII, Indonesia [57], the Museum of Solomos and Eminent Zakynthians [58], the Marble Museum in Carrara, Italy [59]. Temporary exhibitions use gamification for short-term projects, such as the Munch Gallery [60] and Sagamihara City Museum [61].

Meanwhile, physical exhibitions are also included in Table 4. Virtual exhibitions use digital technology to present exhibition content in a virtual environment. Combined with gamification, they provide new access to cultural heritage and museums. For example, digitizing museums with game engines [62]. In addition, add-on games refer to separate game sections added to existing exhibitions to enhance the learning experience [11], such as the Nikos Kazantzakis Muse-um (NKM) [63]; Original games, such as the North Sea Oceanarium [64], apply gamification directly to the design, making gamification part of the show and providing greater immersion [11]. The gamification projects in laboratory settings are mainly in the stage of research and testing, and have not yet been implemented in real exhibition environment. For example, REVAthens [65], Questo [66].

Table 4. Presentation Formats.

Category	Exhibition/museum	Authors
Temporary exhibition	The Munch gallery	[60]
Temporary exhibition	Sagamihara City Museum	[61]
Laboratory project	The 1821 Greek Revolution	[65]
Permanent exhibition	PP-IPTEK Museum TMII, Indonesia	[57]
Permanent exhibition	The North Sea Oceanarium	[64]
Laboratory project		[66]
Permanent exhibition	The Marble Museum in Carrara, Italy	[59]
Temporary exhibition	The Nikos Kazantzakis Muse-um(NKM)	[63]
Virtual exhibition		[62]
Permanent exhibition	The Museum of Solomos and Eminent Zakynthians	[58]

5 Assessment Methods

This section is a collection of assessment methods for applications and projects currently used or developed in gamification exhibitions. These evaluation methods cover many aspects, such as questionnaires, interviews, user experience analysis, thinking aloud, and comprehensively measure the impact of gamification elements, projects and programs on the objectives and contents of the exhibition.

First, the questionnaire is one of the most commonly used evaluation methods in current gamified exhibitions, as shown in Table 5, which is mainly used to measure audience

experience and exhibition effectiveness. Researchers often use SUS questionnaires and Likert scale as criteria and design questions based on these criteria to obtain audience feedback. This can also provide data support for optimization of gamified applications or projects [67, 68].

In addition to a questionnaire survey, the evaluation of gamified exhibits employed methods such as interviews [69], follow-up observations [26], data collection [52], think aloud protocol and reminders [51], State Flow Scale-2 [70].

Table 5. Assessment Methods.

Category	Subject of assessment	Authors
Questionnaire	Collaboration in Interactive Virtual Spaces	[67]
Think Aloud protocol	Program Issues	[51]
Data collection	Gender-specific experiences of use	[52]
Interview	Assessment of the scope of applications	[69]
Follow-up observation	The connection between artifacts and viewers	[26]
State Flow scale-2	The effectiveness of gamification	[70]
Questionnaire	The use of games as an online learning tool	[68]

6 Conclusion

This paper reviews the current development status and use of gamified exhibitions in terms of application domains and scenarios, gamified exhibition design methods and technologies, presentation forms, and assessment methods. This paper helps to clarify the application logic and design framework of gamification in exhibitions, and provides future curators with ideas and references for using gamification as a means of exhibition. In the future, gamified exhibition is expected to provide diversified ideas for exhibition design with the help of emerging technology development, bring a new era of display methods for cultural heritage and exhibition goals, and provide a more innovative exhibition experience for the audience.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. McLean, K.: Museum Exhibitions and the Dynamics of Dialogue. The MIT Press on behalf of American Academy of Arts&Sciences, Cambridge (1999)
2. Kelly, L.: Visitors and Learners: Adult Museum Visitors' Learning Identities Current theories of learning focus on the meanings an individual makes based on their experience-alone, within a social context and as part of a community (Falk and. Prentice (2000)

3. Coenen, T., Mostmans, L., Naessens, K.: MuseUs. J. Comput. Cult. Herit. **6**, 1–19 (2013). https://doi.org/10.1145/2460376.2460379
4. Xu, N., Li, Y., Wei, X., Xie, L., Yu, L., Liang, H.-N.: CubeMuseum AR: a tangible augmented reality interface for cultural heritage learning and museum gifting. Int. J. Hum.-Comput. Interact. **40**, 1409–1437 (2023)
5. He, H., Li, Z., Cheng, X., Wu, J.: Gamified participatory museum experience for future museums. Design, User Experience, and Usability. Application Domains. HCII 2019. NCS, vol. 11585, 195–208. Springer, Cham (2019)
6. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15 (2011). https://doi.org/10.1145/2181037.2181040
7. Chernbumroong, S., Ariya, P., Yolthasart, S., Wongwan, N., Intawong, K., Puritat, K.: Comparing the impact of non-gamified and gamified virtual reality in digital twin virtual museum environments: a case study of Wieng Yong House Museum, Thailand. Heritage **7**, 1870–1892 (2024). https://doi.org/10.3390/heritage7040089
8. Tran, L.-C., et al.: Enhancing the Museum Experience of an Augmented Reality (AR) Art Exhibition Through Digital Exhibit Labels and Gamification. In: SIGHCI 2022 Proceedings Special Interest Group on Human-Computer Interaction, vol. 12 (2022)
9. Araújo, M.B., Alagador, D., Cabeza, M., Nogués-Bravo, D., Thuiller, W.: Climate change threatens European conservation areas. Ecol. Lett. **14**, 484–492 (2011)
10. Morard, S., Sanchez, E., Bonnat, C.: Museum games and personal epistemology: a study on students' critical thinking with a mixed reality game. Int. J. Serious Games **10**, 131–151 (2023). https://doi.org/10.17083/ijsg.v10i4.695
11. Madsen, K.M.: Aalborg Universitet the Gamified Museum Aacritical literature review and discussion of gamification in museums (2020)
12. Trunfio, M., Campana, S., Magnelli, A.: The Impact of Mixed Reality on Visitors' Experience in Museum. "The Ara as It Was" Project in Rome. Progress in IS, pp. 313–322 (2020)
13. Ramić-Brkić, B., Cosovic, M., Rizvic, S.: Cultural Heritage Digitalization in BiH: state-of-the-art review and future trends
14. Thon, S., Serena-Allier, D., Salvetat, C.: Françoise Lacotte: flying a drone in a museum: an augmented-reality cultural serious game in Provence. In: 2013 Digital Heritage International Congress (DigitalHeritage). 28 October 2013–01 November 2013
15. Vlahakis, V., Pliakas, T., Demiris, A., Ioannidis, N.: Design and application of an augmented reality system for continuous, context-sensitive guided tours of indoor and outdoor cultural sites and museums. In: IEEE Virtual Reality Conference, pp. 155–164 (2003)
16. Marques, C.G., Pedro, J.P., Araújo, I.: A systematic literature review of gamification in/for cultural heritage: leveling up, going beyond. Heritage **6**, 5935–5951 (2023)
17. Wang, H., Gao, Z., Zhang, X., Du, J., Xu, Y., Wang, Z.: Gamifying cultural heritage: exploring the potential of immersive virtual exhibitions. Telemat. Inf. Rep. **15**, 100150 (2024). https://doi.org/10.1016/j.teler.2024.100150
18. Ćosović, M., Brkić, B.R.: Game-based learning in museums—cultural heritage applications. Information **11**, 22 (2019). https://doi.org/10.3390/info11010022
19. Paliokas, I., Sylaiou, S.: The use of serious games in museum visits and exhibitions: a systematic mapping study. In: 2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES). Barcelona, Spain, 2016, pp. 1–8 (2016)
20. Falk, E., Liarokapis, F., Peters, C., Mcloughlin, L., Petridis, P., De Freitas, S.: Serious Games in Cultural Heritage Serious Games in Cultural Heritage (2009)
21. Leitão, M.: Design of Immersive Museum Experiences: Improving Visitor Engagement through Co-Creation, Augmented Reality and Gamification (2024)

22. O'Connor, S., Colreavy-Donnelly, S., Dunwell, I.: Fostering engagement with cultural heritage through immersive VR and gamification. In: Liarokapis, F., Voulodimos, A., Doulamis, N., Doulamis, A. (eds.) Visual Computing for Cultural Heritage. Springer Series on Cultural Computing, LNCS, pp. 301–321. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37191-3_16
23. Lien, Y.-N.: Mobile edutainment system -- national palace museum as an example. In: 2015 3rd International Conference on Future Internet of Things and Cloud, pp. 597–601 (2015)
24. Houtkamp, J., van de Kraats, R., Bulens, J., Bulens, J.: SoilGolf: a mixed reality game for the World Soil Museum (2014)
25. Cesário, V.: Guidelines for combining storytelling and gamification. In: 2019 CHI Conference on Human Factors in Computing Systems. CHI EA '19 (2019)
26. Coelho, A., Costa, L.M.: The integration of augmented reality and the concept of sticker album collection for informal learning in museums. Communications in Computer and Information Science, vol. 725, pp. 107–115. Springer, Cham (2017)
27. Yiannoutsou, N., et al.: On establishing contact with cultural objects: the role of a location based game in supporting visitors to engage with contemporary art. In: Proceedings of the 8th International Conference on Interactive Mobile Communication Technologies and Learning. IMCL2014, pp. 290–295 (2014)
28. Moreno Sanchez, I., Barinaga, B., Martin, G., Gutierrez, A., Contreras, R.V.E., Navarro-Newball, A.A.: Archaeological simulation to explore the Chimú Culture from Peru. In: 2019 International Conference on Virtual Reality and Visualization (ICVRV), pp. 266–267 (2019)
29. Xipnitou, M., et al.: From discovery to exhibition - recomposing history: digitizing a cultural educational program using 3D modeling and gamification. Communications in Computer and Information Science, pp. 337–349 (2019)
30. Kontogianni, G., Georgopoulos, A.: A realistic Gamification attempt for the Ancient Agora of Athens. 2015 Digital Heritage, Granada, Spain, 2015, pp. 377–380 (2015)
31. Sousa, F., Nunes, J., Santos, C., Magalhães, J., Vairinhos, M.: MusA: a gamified virtual reality museum. In: GAME ON 2021. Aveiro, Portugal (2021)
32. Coelho, A., van Zeller, M., Cardoso, P., Santos, L., Vaz, R., Raimundo, J.: Gamifying the museological experience. In: XCR, pp. 5–8 (2020)
33. Koutsabasis, P., Vosinakis, S.: Kinesthetic interactions in museums: conveying cultural heritage by making use of ancient tools and (re-) constructing artworks. Virtual Reality **22**, 103–118 (2017). https://doi.org/10.1007/s10055-017-0325-0
34. Cesário, V., Coelho, A., Nisi, V.: This is nice but that is childish. Rapid Communications Presentation. CHI PLAY'19, October 22–25, 2019, Barcelona, Spain, pp. 159–169 (2019)
35. Stavropoulou, G., et al.: Novel museum digitalization framework: the use case of athens museum of paleontology and geology. In: 2023 14th International Conference on Information, Intelligence, Systems & Applications, pp. 1–7 (2023)
36. Grammenos, D., et al.: Macedonia from fragments to pixels: a permanent exhibition of interactive systems at the archaeological Museum of Thessaloniki. In: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F., Caffo, R. (eds.) Progress in Cultural Heritage Preservation. EuroMed 2012. LNCS, vol. 7616, pp. 602–609. Springer, Berlin, Heidelberg. (2012). https://doi.org/10.1007/978-3-642-34234-9_62
37. Ceccacci, S., Generosi, A., Leopardi, A., Mengoni, M., Mandorli, A.F.: The role of haptic feedback and gamification in virtual museum systems. J. Comput. Cult. Herit. **14**, 1–14 (2021). https://doi.org/10.1145/3453074
38. Li, Z., Zhang, Q., Xu, J., Li, C., Yang, X.: Gamification of virtual museum curation: a case study of Chinese bronze wares. Herit. Sci. **12** (2024)
39. Hammady, R., Ma, M., Temple, N.: Augmented reality and gamification in heritage museums. Serious Games 181–187 (2016). https://doi.org/10.1007/978-3-319-45841-0_17

40. Aguirrezabal, P., Peral, R., Pérez, A.S., Sillaurren, S.: Designing history learning games for museums. In: Virtual Reality International Conference. Article 6, pp. 1–4 (2014)
41. Rubino, I., Barberis, C., Xhembulla, J., Malnati, G.: Integrating a location-based mobile game in the museum visit. J. Comput. Cult. Herit. **8**, 1–18 (2015)
42. Zhang, Y., et al.: Serious game design of cultural heritage education based on the experiential learning cycle model. In: 2021 2nd International Conference on Information Science and Education, pp. 1193–1200 (2021)
43. Rowe, J.P., Lobene, E.V., Mott, B.W., Lester, J.C.: Serious games go informal: a museum-centric perspective on intelligent game-based learning. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) Intelligent Tutoring Systems. ITS 2014. LNCS, vol. 8474, pp. 410–415. Springer, Cham (2014)
44. Antoniou, A., Lepouras, G., Bampatzia, S., Almpanoudi, H.: An approach for serious game development for cultural heritage. J. Comput. Cult. Herit. **6**, 1–19 (2013). https://doi.org/10.1145/2532630.2532633vainli
45. Varinlioglu, G., Halici, S.M.: Gamification of Heritage through Augmented Reality. Blucher Design Proceedings, pp. 513–518 (2019)
46. Noreikis, M., Savela, N., Kaakinen, M., Xiao, Y., Oksanen, A.: Effects of gamified augmented reality in public spaces. IEEE Access. **7**, 148108–148118 (2019)
47. Hammady, R., Ma, M., Strathern, C., Mohamad, M.: Design and development of a spatial mixed reality touring guide to the Egyptian museum. Multimed. Tools Appl. **79**, 3465–3494 (2019). https://doi.org/10.1007/s11042-019-08026-w
48. Matthys, M., De Cock, L., Vermaut, J., Van de Weghe, N., De Maeyer, P.: An "animated spatial time machine" in co-creation: reconstructing history using gamification integrated into 3D city modelling, 4D web and transmedia storytelling. ISPRS Int. J. Geo Inf. **10**, 460 (2021). https://doi.org/10.3390/ijgi10070460
49. Andritsou, G., Akrivi Katifori, Kourtis, V., Ioannidis, Y.: Momap - an interactive gamified app for the museum of mineralogy. In: 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games), pp. 1–4 (2018)
50. López-Martínez, A., Carrera, Á., Iglesias, C.A.: Empowering museum experiences applying gamification techniques based on linked data and smart objects. Appl. Sci. **10**, 5419 (2020). https://doi.org/10.3390/app10165419
51. Kontiza, K., Liapis, A., Emma Jones, C.: Reliving the Experience of Visiting a Gallery: Methods for Evaluating Informal Learning in Games for Cultural Heritage (2020)
52. Sprugnoli, R., Guerini, M., Moretti, G., Tonelli, S.: Are these artworks similar? Analysing visitors' judgements on aesthetic perception with a digital game. J. Comput. Cult. Herit. **14**, 1–14 (2021). https://doi.org/10.1145/3461663
53. López-Martínez, A., Iglesias, C.Á., Carrera, Á.: Gamified smart objects for museums based on automatically generated quizzes exploting linked data. In: 2020 16th International Conference on Intelligent Environments (IE). Madrid, Spain, 2020, pp. 132–139 (2020)
54. Donadio, M.G., Principi, F., Ferracani, A., Bertini, M., Bimbo, A.D.: Engaging museum visitors with gamification of body and facial expressions. In: Proceedings of the 30th ACM International Conference on Multimedia, pp. 7000–7002 (2022)
55. Sheidin, J., Ankava, L., Iyov, N., Unkelos, N.: Bringing the sea back to life: an augmented reality voyage in a maritime heritage museum. Advanced Visual Interfaces and Interactions in Cultural Heritage (AVICH 2024) (2024)
56. Nofal, E., Panagiotidou, G., Reffat, R.M., Hameeuw, H., Boschloos, V., Moere, A.V.: Situated tangible gamification of heritage for supporting collaborative learning of young museum visitors. J. Comput. Cult. Herit. **13**, 1–24 (2020)
57. Kristianto, K., Dela, K., Santoso, H.: Implementation of Gamification to Improve Learning in Museum (2018)

58. Yiannoutsou, N., Papadimitriou, I., Komis, V., Avouris, N.: "Playing with" museum exhibits. In: Proceedings of the 8th International Conference on Interaction Design and Children - IDC '09 (2009). https://doi.org/10.1145/1551788.1551837
59. Dini, R., Paternò, F., Santoro, C.: An environment to support multi-user interaction and cooperation for improving museum visits through games. In: Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services (2007)
60. Pierroux, P., Ludvigsen, S.: Communication Interrupted: Textual Practices and Digital Interactives in Art Museums. The Connected Museum: Social Media and Museum Communication (2013)
61. Shirai, A., Kose, Y., Minobe, K., Kimura, T.: Gamification and construction of virtual field museum by using augmented reality game "Ingress." In: Proceedings of the 2015 Virtual Reality International Conference, pp. 1–4 (2015). https://doi.org/10.1145/2806173.2806182
62. Lepouras, G., Vassilakis, C.: Virtual museums for all: employing game technology for edutainment. Virtual Reality **8**, 96–106 (2004). https://doi.org/10.1007/s10055-004-0141-1
63. Apostolellis, P., Bowman, D.A.: Small group learning with games in museums. In: Proceedings of the 14th International Conference on Interaction Design and Children (2015)
64. Vistisen, P., Selvadurai, V., Krishnasamy, R.K.: Applied Gamification in Self-guided Exhibitions: Lessons learned from theory and praxis
65. Gazi, A., et al.: REVAthens: bringing athens of the revolution to life through museum theatre methodology and digital gamification techniques. Communications in computer and information science, pp. 166–183 (2021)
66. Jeon, S.-W., Ryu, G., Moon, S.-J.: Museum gamification design using story elements. Int. J. Adv. Cult. Technol. **8**, 25–32 (2020)
67. Jung, T.H., Tom Dieck, M.C.: Augmented reality, virtual reality and 3D printing for the co-creation of value for the visitor experience at cultural heritage places. J. Place Manag. Dev. **10**, 140–151 (2017)
68. Hong, J.-C., et al.: Using the saliency-based model to design a digital archaeological game to motivate players' intention to visit the digital archives of Taiwan's natural science museum. Comput. Educ. **66**, 74–82 (2013). https://doi.org/10.1016/j.compedu.2013.02.007
69. Sanchez, E.: Palmyre Pierroux: Gamifying the Museum: A Case for Teaching for Games Based Learning (ECGBL 2015) (2015)
70. Weng, Y., Shen, T., Chen, S., Xiao, B.: Gamification in local intangible cultural heritage museums for children: a case design. In: Marcus, A., Wang, W. (eds.) Design, User Experience, and Usability. Application Domains. HCII 2019. LNCS, vol. 11585, pp. 233–245. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23538-3_18

A Framework for the Transmission of Rural Cultural Heritage Through the Integration of Artificial Intelligence and Public Art Curriculum Reform

Jie Ling[1], Junci Xie[1], Ran Wang[1], Li Li[1], Huafang Zhang[1], Cuiyan Zhong[1], Jiaxin Chen[1], and Li Ou-yang[2(✉)]

[1] Zhongkai University of Agriculture and Engineering, Guangzhou 510220, Guangdong, China
[2] The Guangzhou Academy of Fine Arts, Guangzhou 510261, Guangdong, China
oylee@163.com

Abstract. With accelerating globalization and urbanization, rural areas in China face severe issues of population decline and cultural heritage loss. Public art intervention emerges as a promising solution, yet struggles with cultural disconnect and insufficient long-term impact. This study, grounded in constructivist theory and design thinking, explores integrating Artificial Intelligence (AI) into public art education to preserve and innovate rural cultural heritage. Employing interdisciplinary teaching practices and qualitative-quantitative analysis through three rural cases, the research evaluates AI's effectiveness in cultural authenticity, creative efficiency, and implementation feasibility. Results indicate AI tools significantly enhance creative efficiency and cultural expression, but cultural adaptability remains challenging. Future research should further refine educational strategies and technological integration to sustainably advance rural cultural development.

Keywords: Public Art · Artificial Intelligence · Rural Revitalization · Interdisciplinary Education · Cultural Heritage

1 Introduction

1.1 Background

Globalization and urbanization have intensified challenges in rural areas, including population decline, economic stagnation, and cultural erosion. Despite rich cultural heritage and agricultural resources, the outmigration of youth weakens rural vitality. To counter this, the Chinese government has implemented rural revitalization strategies, improving infrastructure, integrating academic expertise, and fostering public art interventions.

However, public art in rural revitalization often lacks sustained impact, with projects limited to temporary exhibitions that fail to yield long-term economic and cultural benefits. Additionally, many initiatives overlook local cultural authenticity, weakening community engagement. Effective public art should not merely decorate rural spaces but actively preserve cultural heritage and address local needs.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 117–127, 2025.
https://doi.org/10.1007/978-3-031-94171-9_10

Despite its potential, public art education has yet to fully integrate rural cultural transmission. Many design programs lack a deep understanding of rural realities, leading to a disconnect between education and practice. Meanwhile, AI technology is revolutionizing creative industries and cultural heritage preservation, accelerating idea generation and design efficiency. Yet, its role in rural public art remains underexplored, raising key questions about how AI can enhance rural-focused public art education.

Taking a public art graduate course at a university in Guangzhou as a case study, this research examines how AI technology can support interdisciplinary learning, rural public art innovation, and material adaptability. The study explores AI's potential to optimize public art design processes while ensuring cultural authenticity and community engagement.

1.2 Research Objectives

This study aims to integrate **AI and interdisciplinary teaching methods** into public art education to enhance rural cultural transmission. The specific objectives are:

1. **Develop an interdisciplinary creative methodology** to bridge knowledge gaps among students from diverse backgrounds.
2. **Ensure AI-assisted design authenticity**, aligning with rural cultural identity and practical environmental demands.
3. **Optimize public art education** by leveraging AI to improve research, ideation, and iterative design efficiency.
4. **Facilitate rural cultural transmission**, integrating AI-generated designs with traditional artistic craftsmanship.

1.3 Research Framework

This study combines constructivist theory, cultural heritage transmission, and the design thinking model, emphasizing the interaction between individual and collective identity. It explores AI's role in generating creative solutions and guiding public art interventions in rural contexts.

The research framework:

1. **Students analyze personal and community cultural needs**, grounding projects in emotional and spatial awareness.
2. **AI is used to conceptualize and refine public art solutions**, enhancing efficiency and cultural relevance.
3. **Traditional craftsmanship integrates with AI**, ensuring practical feasibility and local resonance.
4. **Community engagement and feedback guide design iterations**, fostering meaningful cultural preservation.

This AI-supported public art model bridges traditional and digital artistic practices, introducing new perspectives for rural cultural sustainability.

2 Literature Review

2.1 Constructivist Theory and Public Art Education

Constructivist learning theory posits that knowledge is actively built through experience and contextual interaction [1]. Honebein (1996) outlines seven principles for constructivist education, including learner autonomy, contextualized tasks, collaboration, and reflection [2].

In public art education, this translates to:

- Encouraging students to engage with rural communities through experiential learning.
- Integrating interdisciplinary collaboration, merging art, design, technology, and social sciences.
- Fostering self-directed exploration, allowing students to translate personal narratives into public art.

By structuring education around constructivist engagement and community-based practice, students develop culturally relevant, sustainable public art interventions.

2.2 Cultural Heritage Transmission and Rural Cultural Revitalization

Rural cultural heritage is central to identity and sustainable development. China's Comprehensive Rural Revitalization Plan (2024–2027) emphasizes cultural influence, heritage preservation, and traditional craft revitalization [3]. A collaborative cultural approach—involving multiple stakeholders—ensures the effective transmission and sustainability of heritage [4, 5].

Art-based rural revitalization, akin to Taiwan's community-building models, has emerged as a strategic method for cultural preservation. For example, Yan Qu's Xucun Commune Project successfully engaged local villagers and international artists to co-create rural transformations [6].

Unlike broad art interventions, public art fosters bidirectional cultural exchange, breaking from unilateral "cultural enlightenment" narratives [7]. It strengthens community pride, promotes cultural expression, and recontextualizes heritage, employing sculpture, video, and participatory arts to counteract cultural erosion.

To ensure meaningful heritage transmission, public art must:

- Innovatively reinterpret rural traditions, making them accessible to contemporary audiences.
- Encourage community co-creation, embedding collective cultural identity.
- Ensure long-term sustainability, integrating art into local economic and social frameworks.

2.3 Design Thinking and Public Art Practice

Design thinking applies structured creative problem-solving to align human needs, technological possibilities, and strategic implementation [8]. Stanford D. school's five-phase model—Empathize, Define, Ideate, Prototype, and Test—guides community-centered artistic solutions [9].

In public art, this approach:

- Engages deeply with rural contexts. —students analyze cultural emotions, challenges, and social needs.
- Frames public art interventions with clear objectives—defining community issues.

- Expands creative possibilities through ideation. —encouraging diverse perspectives.
- Visualizes and refines solutions. —using prototyping and iterative testing.
- Ensures public art aligns with cultural authenticity and practical constraints. Embedding design thinking in public art education fosters collaborative, adaptive, and culturally attuned artistic solutions.

2.4 Artificial Intelligence and Innovation in Public Art Design

In recent years, AI technology has been increasingly integrated into artistic and design fields, demonstrating its potential to revolutionize public art through advanced data analysis and generative creativity. AI's ability to synthesize diverse information sources, facilitate rapid ideation, and produce a broad spectrum of design concepts significantly enhances creative efficiency and output quality. Within public art education, AI is particularly valuable for managing complex creative processes and generating diverse design solutions. For instance, AI-driven image generation algorithms enable rapid extraction and transformation of rural cultural elements, helping students quickly refine design directions and produce culturally relevant visual outputs. Additionally, AI tools contribute to evaluating the feasibility of public art installations, optimizing spatial adaptability, and ensuring alignment with local cultural aesthetics.

Despite its advantages, AI-based public art creation presents several challenges. Some students may struggle with AI-generated imagery due to a lack of familiarity with command input, resulting in suboptimal design outcomes. Furthermore, concerns over cultural displacement and authenticity persist, as AI-generated designs may lack the nuanced contextual awareness intrinsic to human artistic interpretation. Thus, this study underscores the importance of establishing community engagement mechanisms and iterative feedback systems within AI-assisted public art education. By fostering a participatory approach, AI tools can be leveraged not only to enhance efficiency but also to ensure that public art projects maintain cultural integrity and effectively serve local community needs. Through this synergy of AI technology and traditional cultural heritage, public art education can drive meaningful innovation while preserving the authenticity of rural artistic traditions.

3 Research Design

3.1 Research Framework and Overall Design

This study integrates an interdisciplinary education approach, combining constructivist theory and the Double Diamond design thinking model, with AI technology and public art curriculum reform as key entry points. The study follows a five-stage design process: empathize, define, ideate, prototype, and test & iterate.

Empathize. Students explore personal experiences and emotions, using identity perception and cultural expression as starting points to understand rural spatial and community needs.

Define. Identifying specific rural challenges such as cultural amnesia, population decline, and ecological degradation, providing clear design focus.

Ideate. AI tools assist in generating diverse creative sketches and solutions.

Prototype. Concepts are developed into visual models or small-scale prototypes for evaluation.

Test & Iterate. Community and expert feedback refine designs for cultural relevance and feasibility.

This structured methodology integrates AI-assisted design, public art education, and rural cultural revitalization into an adaptive learning process.

3.2 Research Objectives and Questions

This study aims to explore how interdisciplinary teaching and AI applications can enhance public art curriculum and preserve rural cultural heritage. Key objectives include:

- **Developing an interdisciplinary creative methodology** that bridges knowledge gaps among students.
- **Ensuring AI-assisted design authenticity**, aligning with rural cultural identity and environmental demands.
- **Optimizing public art education through AI tools**, improving research, ideation, and iterative design phases.
- **Facilitating innovative cultural transmission**, integrating AI-generated design with traditional craftsmanship.

3.3 Research Methods and Data Collection

To achieve these objectives, the study employs qualitative and quantitative research methods:

Qualitative Research Methods. Focus group discussions, faculty-student dialogues, and classroom observations assess student experiences and community feedback on AI-generated designs.

Creative Process Tracking. Monitoring AI tool use, design thinking, and project evolution.

Design Evaluation Metrics. Assessing AI-generated designs based on efficiency, cultural authenticity, and feasibility.

3.4 Course Implementation and Teaching Strategies

The course is structured into eight implementation steps:

Theoretical Foundation. Public art and AI applications introduction.

From Individual to Collective Theme Selection. Personal experiences connected to rural social issues.

Desk Research and AI-Assisted Exploration. Literature review and AI-assisted concept development.

Field Research & Community Interaction. Rural site investigations and participatory research.

Design Generation & Iteration. AI-generated proposals refined with expert and peer feedback.

Prototype & Material Testing. Developing scaled models with culturally and environmentally suitable materials.

Project Implementation. Full-scale public art installations with community participation.

Final Presentation & Evaluation. Expert, educator, and community assessment of project impact.

4 Case Studies and Practical Analysis

4.1 Case Selection and Background

To evaluate the practical impact of this research, three cases were selected: Zuo Tan Village Rural Public Art Project, Zhu Village Qixi Festival Cultural Project, and Desheng Industrial Community Memory Transmission Project. These cases highlight AI's role in public art education across diverse cultural contexts.

4.2 Case Study 1: Zuo Tan Village Rural Public Art Project

Zuo Tan Village in Guangdong Province, a traditional fishing village, transformed a disused power station into a cultural site. Students, inspired by the local maritime culture and their personal study-abroad experiences, developed an art project themed "Returning Home to Rebuild the Village."

Using Doubao AI, students generated concept images featuring fish motifs, refining them into mosaic artwork. Field research confirmed that AI improved design efficiency and cultural relevance, earning strong community approval (Figs. 1, 2, 3, and 4).

Fig. 1. Site conditions.

Fig. 2. AI-generated fish motif design (Tool: Doubao AI, Designer: Wang Ran).

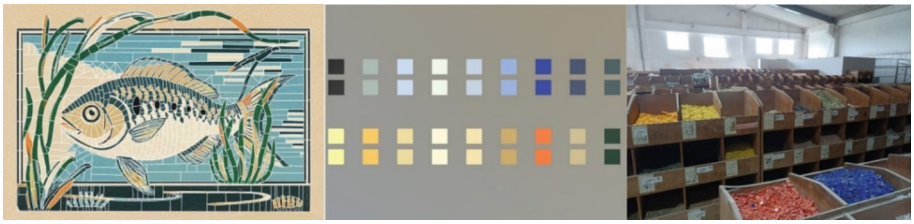


Fig. 3. Finalized design, color selection, and material application.

Fig. 4. Mosaic sample production and spatial integration (Designer: Wang Ran).

4.3 Case Study 2: Zhu Village Qixi Festival Cultural Project

Zhu Village's Qixi Festival, a national intangible cultural heritage, inspired students to explore themes of reunion and love. AI-assisted paper-cut lantern designs, featuring layered cut-outs and internal lighting, won second prize in a provincial design competition. Additional AI-generated archway and bridge designs were proposed but required refinement for authenticity and feasibility (Figs. 5, 6, 7 and 8).

Fig. 5. Zhu Village site documentation (Photographer: Li Li).

Fig. 6. Initial sketches of Qixi Festival elements (Designers: Li Li, Xie Junci, Wang Ran).

Fig. 7. Production process and final execution of the Qixi paper-cut lantern (Designers: Xie Junci, Li Li, Wang Ran).

4.4 Case Study 3: Desheng Industrial Community Memory Transmission

Desheng Industrial Community, a former factory district turned creative hub, retains strong industrial heritage and lion dance culture. Students conducted oral history interviews, using Doubao AI to generate visual artworks and AI-assisted music compositions that capture the community's cultural memory. The AI-music experiment demonstrated new interdisciplinary applications in public art, receiving positive feedback from residents and visitors (Figs. 9, 10 and 11).

Fig. 8. AI-assisted Qixi installation concept (Designer: Li Li).

Fig. 9. Current conditions of Desheng Industrial Park and AI-generated concept sketches (Designer: Xie Junci).

Fig. 10. AI-assisted mosaic artwork and spatial visualization (Designer: Xie Junci).

4.5 Summary

The case studies illustrate how AI enhances public art efficiency, cultural authenticity, and interdisciplinary collaboration. While AI accelerates design processes and diversifies creative output, challenges remain in ensuring cultural fidelity and real-world feasibility. Continued refinement in AI-assisted workflows and community engagement is needed to maximize long-term impact.

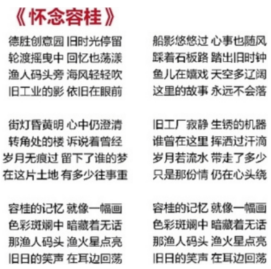


Fig. 11. AI-generated song "Remembering Ronggui" (Composer: Xie Junci).

5 Research Conclusions and Future Directions

5.1 Research Conclusions

This study's case analyses reveal the following key findings:

AI technology enhances design efficiency in public art education, enabling students to rapidly generate diverse creative solutions and supporting interdisciplinary collaboration.

AI aids in the extraction and reinterpretation of rural cultural elements, improving cultural authenticity and community identity in public art projects.

Interdisciplinary teaching strategies strengthen public art education, integrating artistic, technological, and cultural knowledge to foster students' creative and practical skills.

Challenges remain in cultural authenticity and project implementation, requiring further refinement of AI-assisted design workflows and community feedback mechanisms.

5.2 Research Limitations and Future Directions

Research Limitations. Limited case scope: Future studies should examine a broader range of rural communities to enhance generalizability.

Qualitative focus: Incorporating quantitative methods would provide more objective evaluations of AI's impact on public art education.

Future Directions. Develop AI-integrated public art education models, establishing clear pedagogical frameworks and evaluation metrics.

Optimize AI-assisted project implementation, ensuring greater alignment with community needs for sustainable cultural transmission.

Expand interdisciplinary collaboration, exploring AI's role in art, heritage preservation, and creative industries to drive cultural sustainability.

By refining these approaches, future research can further integrate AI into public art education, balancing technological innovation with cultural responsibility in rural heritage.

Acknowledgments. We gratefully acknowledge the financial support from the following programs: Guangdong Provincial First-Class Undergraduate Course (KA24YY044, 6040324137),

2024 Guangdong Provincial Graduate Education Innovation Plan (2024SFKC_054), New Agricultural Science Teaching Research and Reform Practice Project at Zhongkai University of Agriculture and Engineering (PX-44231059), Design Fundamentals Virtual Teaching and Research Office (ZG2023011), Guangzhou Academy of Fine Arts Graduate Demonstration Course Program (60925095), Research on the Innovative Design of Traditional Festival Products in the Guangdong-Hong Kong-Macao Greater Bay Area (20XSB07), the teaching reform project Path Research on Rural Revitalization-Oriented Aesthetic Education Curriculum for 3D Space Based on Knowledge Graphs (60425173), Guangdong Province Second Round "Hundred Million Project" Rural Science and Technology Special Commissioner (KTP20240335), and Research on the Integration of Foshan Intangible Cultural Heritage into Urban Public Art and Digital Innovation (2025-GJ073).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Amineh, R.J., Asl, H.D.: Review of constructivism and social constructivism (2015)
2. Honebein, P.C.: Seven goals for the design of constructivist learning environments
3. The Central Committee of the Communist Party of China and the State Council issue the "Comprehensive Rural Revitalization Plan (2024–2027)". https://www.gov.cn/zhengce/202501/content_7000493.htm. Accessed 16 Mar 2025
4. Song, X., Wang, T.: Cultural revitalization is the "root" and "soul" of rural revitalization — analysis of the importance of rural cultural revitalization and research on the current situation and countermeasures. Shandong Soc. Sci., 176–181 (2019). https://doi.org/10.14112/j.cnki.37-1053/c.2019.04.027
5. Lu, K., Cheng, C.: The changes of public space in traditional villages and the inheritance of rural culture: taking three villages in Zhejiang as examples. J. Guangxi Minzu Univ. (Philosophy and Social Science Edition). **38**, 22–29 (2016)
6. Qu, Y., Wang, C.: Xucun, The Chinese site of artistic rural construction. Times Archit., 44–49 (2015). https://doi.org/10.13717/j.cnki.ta.2015.03.010
7. Sun, Z.: The rural practice of public art. Public Art., 32–39 (2019)
8. Tim-Brown-Design-Thinking. https://designthinkingmeite.web.unc.edu/wp-content/uploads/sites/22337/2020/02/Tim-Brown-Design-Thinking.pdf
9. An Introduction to Design Thinking Process Guide. https://web.stanford.edu/~mshanks/MichaelShanks/files/509554.pdf

Research on the Design of Cultural and Creative APP for Intangible Cultural Heritage Museums from the Perspective of Personalized Customization

Yutong Song(✉), Yichen Li, and Yue Zeng

Beijing City University, No. 269 Bei Si Huan Zhong Lu, Hai Dian District, Beijing, China
2042571698@qq.com

Abstract. Intangible cultural heritage (ICH) is a crucial part of China's cultural identity. However, ICH museums face issues like limited scale, low visitor engagement, and operational challenges. This study proposes a new framework for designing a personalized cultural and creative APP. By using a user-center design approach and iterative prototyping methodology, it explores the appearance, content, and functionality design requirements of the APP. The framework aims to solve problems such as design homogenization and low sales of ICH museum cultural and creative products, and also points out future research directions.

Keywords: Intangible cultural heritage museums · Cultural and creative APP · Personalized customization

1 Background

In the current era of rapid digital development, the significance of intangible cultural heritage (ICH) in shaping a nation's cultural identity and heritage cannot be overstated. ICH represents the living cultural traditions passed down through generations, encompassing a wide range of elements such as oral traditions, performing arts, social practices, and traditional craftsmanship. For China, with its long - standing history and rich cultural tapestry, ICH serves as a crucial link to its past, offering profound insights into the values, beliefs, and creativity of its ancestors.

However, ICH museums, which are tasked with the preservation, exhibition, and promotion of these invaluable cultural assets, face numerous challenges. Their limited scale often restricts the scope of exhibitions, making it difficult to fully showcase the diversity and depth of ICH. Moreover, traditional operation models rely heavily on physical exhibits and in - person visits, resulting in low visitor engagement, especially among younger generations who are more accustomed to digital - first experiences. This not only hampers the dissemination of ICH knowledge but also poses threats to the long - term sustainability of these museums in a competitive cultural landscape.

In the realm of museum - related research, the focus on cultural and creative products has predominantly been on physical items in the past. While physical products have

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 128–143, 2025.
https://doi.org/10.1007/978-3-031-94171-9_11

their merits, the digital domain offers vast untapped potential for ICH museums. Digital platforms can break down geographical barriers, reach a global audience, and provide interactive and immersive experiences that are difficult to achieve through traditional means. Personalized customization, in particular, has emerged as a powerful strategy across various industries, enabling businesses to meet the unique needs and preferences of individual customers. In the context of ICH museums, personalized customization can revolutionize the way visitors interact with cultural heritage, enhancing their connection to ICH and potentially increasing museum revenue through the sale of customized cultural products.

Despite the clear potential, there is a significant lack of research exploring digital platforms and personalized customization for ICH museums. This gap in knowledge has left museums ill-equippedo leverage digital technologies effectively. As a result, many ICH museum cultural and creative products suffer from design homogenization, failing to stand out in the market and attract consumers. The low sales figures further exacerbate the financial challenges faced by these museums, limiting their ability to invest in conservation efforts, research, and innovative exhibitions.

2 Research Context and Concepts

In the current digital - centric era, the landscape of cultural heritage preservation and dissemination is undergoing a profound transformation. Intangible cultural heritage (ICH) lies at the heart of this evolution, presenting both unique opportunities and significant challenges.

ICH is a rich tapestry of cultural expressions passed down through generations. It includes oral traditions, performing arts, social practices, traditional knowledge, and handicrafts. These elements are not only repositories of historical knowledge but also reflect the identity, values, and creativity of communities. For example, the Maori haka in New Zealand is a traditional dance form that embodies the Maori people's warrior spirit, cultural pride, and sense of unity. In India, the intricate art of Madhubani painting, with its symbolic motifs and vibrant colors, tells stories of local mythology, nature, and social life.

However, the institutions responsible for safeguarding and promoting ICH, such as ICH museums, face multiple hurdles. Their physical infrastructure is often limited, constraining the scale and variety of exhibitions. This restricts the ability to fully showcase the diversity of ICH. Traditional operation models, which rely on in - person visits and static displays, struggle to attract and engage modern audiences. The rise of digital media has changed how people access and interact with cultural content. Younger generations, in particular, are more accustomed to digital platforms for learning and entertainment. ICH museums need to adapt to these changing preferences to remain relevant.

2.1 The Transformation of Museums Under the Digital Revolution

The museum sector as a whole is experiencing a digital revolution. Digital technologies offer new ways to overcome geographical barriers and reach a global audience. Virtual

museum tours, for instance, allow people around the world to explore museum collections without leaving their homes. But ICH museums have been slower to embrace digital transformation compared to their mainstream counterparts. There is a dearth of research and development focused on creating digital experiences that are both engaging and respectful of ICH.

Personalized customization is a crucial concept in this research. It involves tailoring products or services to meet the individual needs and preferences of users. In the context of ICH museum apps, personalized customization could mean users being able to choose the content they want to see based on their interests in specific ICH forms. For example, a user interested in traditional Chinese calligraphy could receive customized content about calligraphy techniques, famous calligraphers, and related cultural stories. This approach enhances user satisfaction and engagement by providing a more relevant and unique experience.

User-centered design (UCD) is another important concept. It emphasizes understanding users' needs, expectations, and behaviors throughout the design process. When designing an ICH museum app, UCD would involve conducting user research through surveys, interviews, and usability testing. Designers would then use this information to create an app that is intuitive, easy to use, and meets the diverse needs of different user groups. For example, if user research reveals that many users prefer short, engaging videos over long text - based content, the app can be designed to incorporate more video elements.

Iterative prototyping methodology is a process of creating, testing, and refining prototypes in a cyclic manner. In app design, a basic prototype is first developed, which includes key features such as navigation, content presentation, and interaction elements. This prototype is then tested with users, and feedback is collected. Based on this feedback, the prototype is improved, and the testing - improvement cycle is repeated until the app meets the desired quality and usability standards. This ensures that the final app is user - friendly and effective in achieving its goals.

2.2 The Digital Divide in Intangible Cultural Heritage Preservation

In the contemporary digital age, the digital divide in intangible cultural heritage (ICH) preservation has emerged as a significant issue, creating a disparity in how different cultural heritage institutions manage and promote ICH. This divide is characterized by the gap between mainstream museums, which are often well - equipped with digital resources, and ICH museums, which struggle to keep up with digital advancements.

Mainstream museums typically have access to substantial financial resources, enabling them to invest in state - of - the - art digital technologies. They can afford to digitize their collections comprehensively, creating high - resolution images, 3D models, and interactive virtual exhibits. For example, large - scale art museums in major cities can use advanced scanning techniques to capture every detail of a painting or sculpture, allowing users around the world to view and study these artworks online. These museums also often have dedicated digital teams, including developers, designers, and content creators, who work to enhance the digital user experience.

In contrast, ICH museums face numerous challenges in the digital realm. Limited funding is a major constraint. Many ICH museums operate on tight budgets, with a

significant portion of their resources allocated to basic preservation and day - to - day operations. As a result, they have less money available for digital initiatives. This lack of financial support means that they may be unable to digitize their collections effectively, or they may use outdated digital technologies that provide a subpar user experience.

Another contributing factor to the digital divide is the shortage of digital skills among ICH museum staff. ICH museums often have a staff with expertise in traditional cultural heritage fields, such as anthropology, history, or folklore. However, they may lack individuals with specialized digital skills, like web development, app design, or digital content management. Without the necessary skills, it becomes difficult for these museums to create engaging digital platforms, manage digital assets properly, or keep up with the latest digital trends.

The digital divide has far - reaching implications for ICH preservation. It restricts the accessibility of ICH to a global audience. While mainstream museums can reach millions of users through their digital platforms, ICH museums may struggle to attract a significant online following. This limited reach not only affects the dissemination of ICH knowledge but also reduces the potential for public engagement and support.

Furthermore, in the long - term, the digital divide can undermine the preservation efforts of ICH. Digital technologies play a crucial role in documenting and archiving ICH. Without proper digital infrastructure and skills, ICH museums may be at risk of losing valuable cultural heritage data. For instance, if traditional cultural practices are not digitized in a timely manner, they may be lost as the number of practitioners declines.

2.3 The Role of Personalized Experiences in Cultural Engagement

In an era inundated with information and diverse forms of entertainment, capturing and maintaining individuals' interest in cultural heritage, especially intangible cultural heritage (ICH), has become a significant challenge. Personalized experiences play a crucial role in addressing this challenge, acting as a powerful catalyst for deeper cultural engagement.

1. Tailoring Content to Individual Interests

   One of the key aspects of personalized experiences is the ability to customize content according to users' unique interests. In the context of ICH, this means that instead of presenting a one - size - fits - all approach to cultural information, platforms can offer tailored content. For example, a person interested in traditional Japanese textiles such as kimono - making can be presented with in - depth articles, videos, and interactive exhibits focused on the history, techniques, and symbolism of kimono production. This targeted content not only aligns with the user's existing interests but also has the potential to deepen their knowledge and appreciation of this specific aspect of ICH. By using data analytics and user - inputted preferences, cultural platforms can create personalized content recommendations. This is similar to how streaming services suggest movies or music based on a user's past viewing or listening history. In the realm of ICH, if a user frequently engages with content related to African oral storytelling, the platform can recommend other related stories, or introduce them to the art of storytelling in other cultures with similar themes. [1] This personalized curation of content encourages users to explore ICH more comprehensively, as they are more likely to engage with material that resonates with their pre - existing interests.

2. Enhancing Emotional Connection

   Personalized experiences also have the power to forge a stronger emotional bond between individuals and ICH. When users encounter cultural content that is relevant to their personal lives, hobbies, or values, they are more likely to feel a sense of connection. For instance, someone who practices yoga might feel a special affinity for the philosophical and spiritual aspects of Indian ICH, such as the ancient texts and rituals associated with yoga's origins. A personalized experience could highlight these connections, perhaps by presenting modern - day yoga practices in the context of their historical and cultural roots. This emotional connection is essential for long - term cultural engagement. It transforms passive consumption of cultural information into active participation. Users who feel emotionally invested in ICH are more likely to share their experiences with others, participate in cultural events, and even contribute to the preservation of ICH. They become ambassadors for the cultural heritage, spreading awareness and promoting its continued relevance in contemporary society.
3. Facilitating Active Participation

   Personalized experiences often involve elements that encourage active participation rather than passive consumption. In the context of ICH, this could take the form of interactive workshops, virtual try - ons of traditional costumes, or opportunities to create and share personalized interpretations of cultural artifacts. For example, an ICH - themed app might offer a virtual embroidery workshop where users can learn traditional embroidery techniques from different cultures and create their own digital embroidery patterns. These hands - on experiences are particularly effective in engaging younger generations, who are accustomed to interactive and immersive digital content. By allowing users to actively engage with ICH, personalized experiences not only make learning about cultural heritage more enjoyable but also give individuals a sense of ownership over the cultural knowledge they acquire. This sense of ownership further strengthens their connection to ICH and motivates them to engage with it on an ongoing basis. In conclusion, personalized experiences are a vital component in promoting cultural engagement with ICH. By tailoring content, enhancing emotional connections, and facilitating active participation, they offer a more effective way to introduce people to the rich and diverse world of intangible cultural heritage, ensuring its preservation and continued relevance in the modern era.

2.4 User - Centered Design: A Catalyst for ICH Museum App Success

In the quest to create impactful digital platforms for intangible cultural heritage (ICH) museums, user - centered design (UCD) has emerged as a fundamental approach. UCD places users at the heart of the design process, ensuring that the resulting app meets their needs, expectations, and preferences. This approach is not just a design philosophy but a crucial factor in the success of ICH museum apps, enabling them to effectively engage users and preserve and promote ICH.

1. Understanding User Needs

   The first step in UCD for ICH museum apps is to gain a deep understanding of user needs. Different user groups, such as students, researchers, tourists, and local community members, have diverse interests and requirements. For students, the app might serve as an educational tool, offering interactive lessons, quizzes, and access to

primary sources related to ICH. Researchers, on the other hand, may seek in - depth academic resources, detailed documentation of cultural practices, and the ability to collaborate with other scholars.

To gather this information, designers employ various research methods. Surveys can be distributed to a wide range of potential users to identify common interests, pain points, and expectations. For example, a survey might reveal that many users struggle to find accurate translations of ICH - related terms in multiple languages. Interviews, both in - person and online, allow for more in - depth exploration of user experiences and motivations. Additionally, usability testing of existing museum digital resources can uncover usability issues, such as complex navigation or unresponsive interfaces.

2. Creating an Intuitive User Interface

Based on the insights gained from user research, designers can create an intuitive user interface (UI) for the ICH museum app. The UI should be easy to navigate, with clear labels and logical menu structures. For instance, instead of overwhelming users with a jumble of information, the app can use a modular design, where different sections are dedicated to specific ICH themes, such as traditional music, handicrafts, or performing arts.

Visual design also plays a crucial role. The app's color scheme, typography, and imagery should be inspired by ICH elements while maintaining a modern and accessible look. For example, if the app focuses on Chinese ICH, it could incorporate colors from traditional Chinese art, like the red and gold associated with good luck and prosperity. Icons and graphics should be easily recognizable and relevant to the content, helping users quickly identify different functions and sections of the app.

3. Ensuring Inclusivity

Inclusivity is a key aspect of UCD for ICH museum apps. The app should be accessible to users with disabilities, including those with visual, auditory, or motor impairments. This can be achieved through features such as screen reader compatibility, closed - captioning for videos, and easy - to - use touch controls. For example, providing alternative text descriptions for images ensures that visually impaired users can understand the content, while adjustable font sizes and high - contrast color options accommodate users with visual difficulties.

Moreover, the app should be culturally inclusive, respecting and representing the diversity of ICH. It should avoid cultural stereotypes and present ICH in an accurate and respectful manner. This means involving cultural experts and community members in the design process to ensure that the app's content and representation are authentic.

4. Iterative Improvement

UCD is an iterative process. Once the initial app prototype is developed, it is tested with users, and feedback is collected. This feedback can highlight areas for improvement, such as slow loading times, unclear instructions, or lack of certain features. Based on this feedback, designers make adjustments to the app, creating a new version for further testing.

For example, if users complain that the app's search function is ineffective, designers can improve the search algorithm or add more search filters. This iterative cycle of testing and improvement continues until the app meets the needs and expectations

of users. By continuously refining the app based on user feedback, ICH museum apps can stay relevant and engaging in a rapidly evolving digital landscape.

In conclusion, user - centered design is a powerful catalyst for the success of ICH museum apps. By understanding user needs, creating an intuitive UI, ensuring inclusivity, and engaging in iterative improvement, these apps can effectively connect users with ICH, fostering a deeper appreciation and understanding of this valuable cultural heritage.

2.5 Iterative Prototyping: Refining ICH Museum Apps for Optimal Performance

In the development of intangible cultural heritage (ICH) museum apps, iterative prototyping stands as a cornerstone for achieving optimal performance. This methodology involves a cyclical process of creating, testing, and refining prototypes, ensuring that the final app meets the diverse needs of users while effectively presenting ICH.

1. The Initial Prototype: Laying the Foundation

   The journey of iterative prototyping begins with the creation of an initial prototype. This prototype serves as a basic framework, incorporating the fundamental features and functions that the app aims to offer. For an ICH museum app, this might include a simple navigation system, basic content presentation such as text - based descriptions of ICH items, and a few interactive elements like buttons for accessing different sections.

   During this stage, developers draw on the insights gathered from user - centered design research. They identify the core user needs and translate them into tangible app features. For example, if user research indicates that visitors are interested in learning about the history and techniques behind traditional ICH crafts, the initial prototype might include a section dedicated to these topics, with brief written explanations and low - resolution images.

   However, the initial prototype is not expected to be perfect. It is a starting point, a rough draft that allows developers to visualize the app's basic structure and functionality. This early version may have limitations in terms of design polish, performance, and the depth of content, but it provides a platform for further development.
2. Testing the Prototype: Gathering User Feedback

   Once the initial prototype is ready, the next crucial step is testing. User testing is at the heart of iterative prototyping, as it provides valuable insights into how real - world users interact with the app. There are various methods for conducting user tests, such as usability testing, where users are asked to perform specific tasks within the app, like finding a particular ICH item or using a customization feature.

   During usability testing, developers observe users' actions, note any difficulties they encounter, and listen to their feedback. For instance, users might struggle to find the search bar, or they might be confused by the way content is organized. This feedback is invaluable as it highlights areas that need improvement.

   Another form of testing is user acceptance testing, where a broader group of potential users evaluate the app's overall appeal, functionality, and whether it meets their expectations. They might provide feedback on aspects such as the app's visual design, the relevance of the content, or the ease of sharing information from the app.

3. Refinement Based on Feedback: Enhancing the Prototype

   After collecting user feedback, developers embark on the refinement phase. They analyze the feedback to identify common issues and areas for improvement. For example, if multiple users complain about the slow loading time of images in the app, developers can optimize the image compression or improve the app's caching mechanism.

   Based on these findings, developers make targeted changes to the prototype. This could involve modifying the user interface design to make it more intuitive, adding or improving features to enhance functionality, or updating the content to make it more engaging and accurate. For instance, if users suggest adding more multimedia elements like videos or audio recordings to the ICH craft section, developers can incorporate these elements into the next version of the prototype.

   The refined prototype is then tested again, and the cycle continues. Each iteration builds on the previous one, gradually eliminating bugs, improving performance, and enhancing the user experience. This iterative process ensures that the app evolves in response to user needs and preferences, rather than being a static product developed in isolation.

4. The Impact on ICH Museum App Performance

   Iterative prototyping has a profound impact on the performance of ICH museum apps. By continuously refining the app based on user feedback, developers can create an app that is highly user - friendly. A well - designed and easy - to - use app is more likely to attract and retain users, increasing user engagement with ICH.

   In terms of functionality, iterative prototyping allows for the addition of features that enhance the app's utility. For example, features like personalized recommendations, augmented reality experiences, or social sharing capabilities can be added and refined over multiple iterations. These features not only make the app more appealing but also contribute to a more immersive and interactive exploration of ICH.

Moreover, iterative prototyping helps ensure that the app effectively conveys the value and significance of ICH. Through successive refinements, the content can be made more accurate, engaging, and accessible, facilitating a deeper understanding of ICH among users. In summary, iterative prototyping is an essential process for creating ICH museum apps that are optimized for performance, user satisfaction, and the preservation and promotion of ICH.

3 Research Context and Concepts

Based on the characteristics of intangible cultural heritage museums, the composition of personalized customized cultural and creative elements is more complex. In order to better integrate elements and achieve innovation in museum cultural and creative industries today, personalized customization of cultural and creative products requires a creative orientation based on cultural and creative appearance, content, and function, designing more products aimed at enhancing the attractiveness and added value of non heritage cultural and creative products, meeting the personalized needs of consumers, and promoting the inheritance and dissemination of intangible cultural heritage.

Intangible cultural heritage has unique colors, patterns, and craftsmanship, with strong regional characteristics and cultural charm. When designing non heritage cultural and creative products, it is necessary to integrate the artistic elements and creative techniques of intangible cultural heritage into modern product design concepts, and produce cultural and creative products that meet the aesthetic tastes of people today. Non legacy cultural and creative products, like traditional products, need to not only be practical and aesthetically pleasing, but also constantly innovate their fun and fashion appeal in order to better attract consumers, especially modern young people who pursue novelty and trends.

3.1 Personalized Customization Method

1. Function driven

   Firstly, determine the main functional requirements of cultural and creative products, such as practical functions, decorative functions, and commemorative functions. Based on different target groups and usage scenarios, clarify functional positioning. In terms of practical functionality, for personalized customized products, it is necessary to expand product categories and maximize user choices to meet their selectivity and sense of participation. From the perspective of user life scenarios, the purpose is to meet users' needs for purchasing cultural and creative products to enrich their life and culture [2].
2. Content customization

   After determining the functionality, appropriate content was designed around the functionality. The content can include cultural elements, storytelling, etc. And specific cultural connotations, historical stories, traditional skills represented by intangible cultural heritage projects can be selected as content themes. For example, cultural and creative products with the theme of Paper Cuttings art can show the historical origin of Paper Cuttings, the specific styles of Paper Cuttings in different regions, and the history of Paper Cuttings masters. They can be presented in the form of text, images, sound, or video, allowing users to have a clear understanding of intangible cultural heritage when using the product. This not only enriches the connotation of the product, but also enhances the cultural experience of users.
3. appearance customization

   The appearance design should match the functionality and content, while meeting the aesthetic needs of users. Multiple appearance styles can be provided for users to choose from, such as minimalist style, retro style, cartoon style, etc. Taking ceramic cultural and creative products as an example, if the function is to serve as tea sets, the content revolves around a specific intangible cultural heritage ceramic firing technique. The appearance can be designed in a traditional blue and white porcelain style, a brightly colored modern style, or a cute style with cute patterns to meet the preferences of different users. Customization of appearance can make products visually more attractive and increase users' desire to purchase.

The appearance design should match the functionality and content, while meeting the aesthetic needs of users. Users can choose from minimalist style, retro style, cartoon style, etc. If the function is used as a tea set, then the problem revolves around a specific

technique of firing intangible cultural heritage ceramics. This appearance can be designed in a traditional blue and white porcelain style, a brightly colored modern style, or a cute patterned cute style to meet the choices of different users. This makes the product visually more attractive and increases users' desire to purchase [3].

3.2 Customization of Intangible Cultural Heritage Techniques as Products

1. Provide entity customization toolkit

   Cultural and creative apps can provide users with customized physical toolkits. The toolkits related to specific intangible cultural heritage techniques have authentic tools and materials. For example, for traditional embroidery techniques, there can be various types of embroidery needles, threads, fabrics, etc. in the tool kit; For wood carving techniques, small carving knives, wood, etc. can be provided. Each toolkit includes detailed user manuals and tutorials to assist users in creating using these tools. This toolkit not only provides users with the opportunity to personally experience intangible cultural heritage skills, but also increases the fun and interactivity of the product.
2. User creation process

   After receiving the toolkit, users can use the tools and materials to create intangible cultural heritage skills according to their preferences and creativity. For example, in embroidery intangible cultural heritage creation, users can choose their favorite patterns or design their own patterns, and then use embroidery needles and threads to customize their creations on the embroidery fabric. For wood carving intangible cultural heritage techniques, users can use toolkits to create according to the auxiliary lines on customized wood materials. The user creation process not only meets the personalized needs of users, but also allows them to have a deeper understanding of intangible cultural heritage skills and their sense of identification with intangible cultural heritage.

3.3 Develop Product Modules that Are Compatible with Intangible Cultural Heritage Functions

1. The embodiment and integration of the functional nature of intangible cultural heritage

   The function of intangible cultural heritage is reflected in the perfect combination of practicality and artistry. The use of traditional intangible cultural heritage bamboo weaving techniques, such as bamboo baskets and bamboo baskets, to improve bamboo products has been widely applied in daily life. They not only have practical value for storing items, but also showcase exquisite weaving techniques and unique artistic beauty. In the case of similar functions in modern products, both can be integrated. For example, combining improved bamboo weaving techniques with modern home decor, applying traditional bamboo weaving techniques to lamps and vases, not only preserves the natural and rustic beauty of bamboo weaving, but also meets the personalized needs of modern people for home decoration. Another example is the traditional intangible cultural heritage Paper Cuttings art, which can add strong cultural flavor and artistic charm to book binding, packaging design and other fields by combining its unique decorative and cultural meaning with modern graphic design.

2. Integration into Product Structure Modularization

   In modular design of product structure, integrate product modules that integrate intangible cultural heritage as independent functional units. If it is a home product, the bamboo weaving decoration module can be used as a part of the product and can be detachably connected to the main structure. The advantage of this design is that users can choose different intangible cultural heritage modules to combine according to their preferences and needs, achieving personalized product customization. At the same time, this modular structure also facilitates product maintenance and upgrades, enhancing the sustainability of the product. Taking lighting fixtures as an example, the bamboo lampshade module can be designed as a replaceable structure, allowing users to choose different styles of bamboo lampshades for replacement according to different scenes and moods. Alternatively, in a furniture product, traditional wood carving decorative modules can be added as optional accessories. Users can decide whether to add wood carving decorations based on their home style and personal preferences, thereby achieving product diversification and personalization.
3. Integration of modular design elements

   In the modularization of design elements, the elements of intangible cultural heritage can be disassembled and reorganized to form independent design modules. These modules include different elements such as patterns, colors, and materials, which are combined in various ways to create rich and diverse design effects.

Pattern module: Provides users with various patterns related to intangible cultural heritage, such as traditional embroidery patterns, wood carving patterns, New Year painting patterns, etc. Users can choose to apply it to different positions of cultural and creative products according to their preferences.

Color module: Determine color schemes based on a series of color combinations representing Chinese intangible cultural heritage, such as traditional Chinese red and blue colors, or based on different intangible cultural heritage characteristic projects. Users can choose their favorite colors to customize color combinations for different modules of the product.

Material module: By selecting different materials such as wood, ceramic, fabric, metal, etc. Each material has its unique texture and characteristics, and users can choose the appropriate material according to the product's purpose and personal preferences. They can also combine and assemble materials from different modules.

Shape module: By providing multiple shape options such as circles, squares, polygons, etc., or designing unique shapes based on specific intangible cultural heritage elements, the shape is based on a certain intangible cultural heritage graphic. Users can choose the shape of the product according to their own needs and aesthetics.

Accessory module: Cultural and creative products provide various replaceable accessories to increase the personalization of the product. For example, providing laptops with accessories such as covers of different materials, bookmarks of different types, and pen clips of different colors. Users can choose accessories to combine according to their preferences.

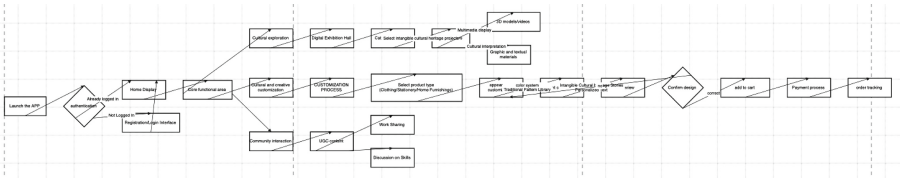
Packaging module: Choose various forms of packaging, such as simple and environmentally friendly packaging, exquisite gift box packaging, etc. Users can choose suitable packaging according to their purchase purpose.

By integrating the product modules of intangible cultural heritage into the modularization of product structure and design elements, a deep integration of intangible cultural heritage and modern product design can be achieved, creating products with unique cultural value and market competitiveness. At the same time, this modular design method also provides new ideas and approaches for the inheritance and innovation of intangible cultural heritage.

3.4 App UI Interaction Design

1. Design principles

   User centered: Gain a deep understanding of the needs, expectations, and usage habits of various users of the Intangible Cultural Heritage Museum app, including students, researchers, tourists, and local community members. For example, students may be more concerned about learning functions and hope to obtain rich knowledge explanations and interactive learning resources through the APP; Researchers require precise and in-depth academic materials. Based on these different needs, design UI interaction interfaces that meet various user operating habits and psychological expectations, allowing users to easily find the required functions and information.



Cultural integration: Cleverly integrating the characteristic elements of intangible cultural heritage into UI interaction design. In terms of color selection, we can learn from the color matching in traditional art, such as the red and green contrast colors commonly used in Chinese Paper Cuttings art, to create a lively and cultural visual effect; In terms of icon design, drawing inspiration from traditional handicraft tools or cultural symbols, such as designing functional icons in the shape of embroidery needles, not only increases cultural recognition but also makes the interface more unique. At the same time, the production process of traditional techniques, such as the dynamic effects of ceramic firing, is showcased through dynamic design, allowing users to experience the charm of intangible cultural heritage during the operation process.

Simplicity and ease of use: Ensure that the interface is concise and clear, avoiding too many complex design elements that interfere with user operations. Adopting a clear layout structure, the content is classified according to different themes and functions, such as traditional skills, performing arts, folk culture, etc., allowing users to quickly locate the content they are interested in. Simplify the operation process, for example, in the customization function, simplify the operation through “one click customization” and other methods to reduce the user’s operating costs. At the same time, provide clear operating instructions and feedback, and provide visual or auditory feedback in a timely manner when the user clicks the button, letting the user know that the operation has been executed (see Fig. 1).

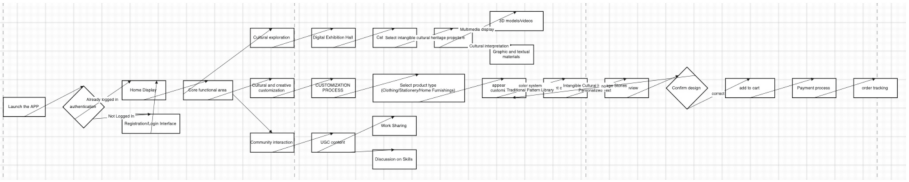


Fig. 1. App usage flowchart

2. Interface layout

Homepage design: The homepage adopts a simple and atmospheric layout, with large and exquisite images of intangible cultural heritage as the background, attracting users' attention. Set up a search bar at the top for users to quickly find specific cultural content; Below, the core functions of the app, such as "Cultural Exploration," "Product Customization," and "Interactive Community," are displayed through concise icons and text. At the same time, personalized recommendation areas are set up to display relevant intangible cultural heritage projects or cultural and creative products based on users' browsing history and interest preferences, guiding users to further explore (see Fig. 2).

Fig. 2. App homepage interface design

Content display page: Adopt appropriate layout according to different content types. For cultural knowledge introduction pages that focus on text, use clear and easy to read fonts, segment them reasonably, and pair them with relevant images or charts to assist understanding; For multimedia content such as videos or audio, set up dedicated playback areas and provide playback control buttons and progress bars. Add bookmark, share, and comment buttons on the page to facilitate users in tagging interesting content, sharing with others, and interacting with others.

Customized Function Page: The customized function page adopts modular design, clearly categorizing and displaying customized options. For example, in the customization of cultural and creative products, modules such as "appearance customization", "content customization", and "function customization" are set up separately. Within each module, a rich variety of choices are provided, such as in appearance customization, where users can choose different style templates and color combinations; In content customization, users can add their favorite cultural elements, stories, etc. At the same time, it provides real-time preview function, allowing users to view the final effect at any time during the customization process, improving user participation and satisfaction.

3. Interaction design

   Navigation interaction: Adopt a simple and intuitive navigation method, such as a bottom navigation bar or sidebar. The bottom navigation bar usually includes main function entrances such as "Home", "Explore", "Customize", "My", etc., making it convenient for users to quickly switch between different function pages; The sidebar can be expanded when needed to display more functional options and content categories, saving page space. During the navigation process, use clear icons and text labels to allow users to clearly understand the function of each navigation item. At the same time, add dynamic effects to the navigation, such as gradient effects when switching pages, to enhance the smoothness of the user experience.

Operation interaction: Optimize the interaction methods of various operations to improve the convenience of user operations. For example, in the search function, automatic association and popular search keyword recommendations are provided to help users find the desired content faster; In customization operations, gesture operations such as drag and drop, zoom, etc. are supported to facilitate users in adjusting the position and size of customized elements. In addition, adding a feedback mechanism for operations, such as uploading a large number of images for customization, displaying an operation progress bar when users perform complex operations, so that users can understand the progress of the operation.

Feedback interaction: Timely provide user operation feedback to enhance the interaction between users and the APP. When a user completes an operation, such as posting a comment or customizing a product, they are notified of the success of the operation through pop ups, vibrations, or beeps; If the operation fails, clearly indicate the reason for the failure and provide corresponding solutions. During the operation of the app, when loading content, display loading animations to avoid user anxiety caused by waiting. At the same time, personalized feedback is provided based on user behavior, such as recommending relevant in-depth content or activity information when users frequently browse a certain type of intangible cultural heritage content.

4 The Advantages and Challenges of Personalized Customized Cultural and Creative Apps for Intangible Cultural Heritage Museums

4.1 Advantage

Meeting Individual requirements: A personalized customized cultural and creative app that can provide diverse product choices and customized solutions based on users' different needs and preferences, offering a variety of product choices and customized solutions to meet users' personalized needs. Users can customize their personal cultural and creative products based on their aesthetics, interests, and usage scenarios to enhance their sense of participation and satisfaction.

Promote the inheritance and dissemination of intangible cultural heritage.

By combining intangible cultural heritage with modern technology and designing and promoting cultural and creative apps, more people can become aware of intangible cultural heritage. In the process of customizing cultural and creative products, users can gain a deeper understanding of the connotation and value of intangible cultural heritage, enhancing their sense of identification and protection of intangible cultural heritage. At the same time, cultural and creative apps also provide a platform for intangible cultural heritage inheritors and museums to showcase and promote intangible cultural heritage, promoting the inheritance and development of intangible cultural heritage.

1. Enhance market competitiveness

   Personalized customized cultural and creative apps can bring new sales channels and business models to intangible cultural heritage museums, and have a strong role in enhancing the market competitiveness of intangible cultural and creative products. Compared with traditional sales methods for cultural and creative products, cultural and creative apps have a wider user base and higher sales efficiency. By utilizing advantages such as personalized customization and modular design, the added value and differentiated competitive advantage of products can be improved, meeting consumers' demand for high-quality and personalized cultural and creative products.

4.2 Challenge

1. Technical implementation difficulty

   The design and development of personalized cultural and creative APP involves technologies in many fields, such as mobile Internet technology, graphics and image technology, database technology, etc. At the same time, user experience and product quality also need to be considered, and technological implementation requires a significant investment of manpower, material resources, and financial resources.
2. Balancing the Protection and Innovation of Intangible Cultural Heritage

   In the design process of personalized customized cultural and creative apps, it is necessary to fully consider the balance between the protection of intangible cultural heritage and innovation. On the one hand, we should respect the traditions and characteristics of intangible cultural heritage and avoid excessive commercialization and entertainment development of it; On the other hand, it is necessary to combine modern technology and design concepts to innovate and develop intangible cultural heritage, making it more in line with the needs and aesthetics of modern society.

3. The uncertainty of user demand

   User needs are uncertain and diverse. Accurately grasping user needs and providing products and services that meet user needs is a challenge that personalized customized cultural and creative apps need to face. Continuously optimize products and services and improve user satisfaction through market research, user feedback, and data analysis.

References

1. Chandra, S., Verma, S., Lim, W.M., Kumar, S., Donthu, N.: Personalization in personalized marketing: trends and ways forward. Psychol. Mark. **39**, 1529–1562 (2022)
2. Ferrer-Yulfo, A.: Transforming museum education through intangible cultural heritage. J. Museum Educ. **47**(3), 319–330 (2022)
3. Tu, J.C., Liu, L.X., Cui, Y.: A study on consumers' preferences for the palace museum's cultural and creative products from the perspective of cultural sustainability. Sustainability **11**(13), 3502 (2019)

Digital Storytelling and Narratives in Cultural Heritage Sites. A Case Study on Interactive Museum Experience Design

Yijia Sun[1,2] and Fangzhou Dong[2,3](✉)

[1] Southeast University, Nanjing, China
ysun0173@student.monash.edu
[2] Suzhou Industrial Park Monash Research Institute of Science and Technology, Suzhou, China
fangzhou.dong@monash.edu
[3] Monash University, Melbourne, Australia

Abstract. This paper explores the integration of interactive digital storytelling in cultural heritage sites to enhance visitor engagement and create meaningful personal narratives. Using a participatory ethnographic approach within a constructive design research framework, the study presents a museum experience design as a case study. Visitors engage with exhibits on-site through a mobile app to collect artifacts and customize souvenirs supported by AI technology, which deconstructs features into four layers: outer contour, inner contour, texture, and base color. These souvenirs serve as gateways to a personal library, allowing visitors to access detailed information about cultural heritage anytime and anywhere. By blending on-site and off-site interactions, the design fosters deeper connections between visitors and cultural heritage. The findings highlight AI deconstruction as a creative method, souvenirs as narrative gateways, and personal narratives as a cohesive element in the museum experience. This research provides insights for leveraging digital storytelling to enhance cultural heritage dissemination and inspire innovation in design for heritage sites.

Keywords: Museum Experience · Digital Storytelling · Narratives · Case Study

1 Introduction

For an extended period, as cultural heritage sites, museums have played an important role in preserving physical artifacts such as historical objects, artworks, and photographs. Concurrently, cultural heritage extends beyond the tangible artifacts we can observe and touch; it also encompasses the intangible elements, including how individuals interact with these sites and their subjective interpretations of the transformations occurring within them [1]. With this broader concept in mind, museums are increasingly demonstrating a commitment to fostering new modes of engagement and participation by providing diverse resources that enrich interactions, deepen visitor involvement, and enhance educational experiences [2]. During this process, narratives allow museums to create more innovative methods to present and enrich the stories behind their exhibits [3], attract visitors' attention and increase their interest through active engagement.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 144–151, 2025.
https://doi.org/10.1007/978-3-031-94171-9_12

In the digital era, heritage sites transcend the role of mere narrators that convey static narratives through curatorial design [4]. Instead, they emerge as co-creators of narratives, dynamically engaging with visitors and shaping stories collaboratively through the interplay of visitor experiences within the space. Although limited efforts have been made to integrate interactive narratives in cultural heritage sites, there are signs that this approach could significantly enhance visitor engagement by creating digital storytelling [5]. Therefore, this article explores how interactive digital storytelling highlights museums' intangible heritage aspects, fostering visitors' sense of place and identity within cultural and historical contexts. Through the MuseCode case study, this article demonstrates how digital storytelling integrates real-time engagement and reflective interactions to create a seamless and immersive journey. By deepening connections to cultural heritage across all phases of the visitor experience, this research highlights the potential of post-visit interactions to enrich and sustain engagement with cultural heritage.

2 Related Works

In the view of Bruner [6], narrative is a vital tool for constructing reality, as we organize human events into narratives and derive meaning from them. The accumulation of narratives shapes culture, history, and tradition, providing the foundation for cultural negotiation and social continuity. Digital storytelling represents a modern narrative form that combines elements of traditional storytelling methods with innovative approaches to self-expression [7]. Extensive research has been conducted on interactive digital storytelling in cultural heritage contexts, particularly focusing on enhancing user engagement. Digital storytelling enhances narratives by making them more immersive and interactive, allowing visitors to actively participate in the storytelling process through multimedia elements, personalized experiences, and interactive exhibits.

The extensive penetration of digital technology, such as social media, has introduced entirely new channels and discourses for museum communication, enabling them to transcend traditional physical space constraints and engage with the public in a more immediate and expansive manner [8]. Furthermore, through technologies such as virtual reality (VR) and augmented reality (AR), museums offer immersive experiences that reshape the public's cognition and information acquisition about cultural heritage [9]. In recent studies, generative artificial intelligence (GenAI) is also applied in design projects to explore approaches to engaging visitors in exhibition. For instance, research [1] fully leveraged the advantages of GenAI to design an exhibition that enables visitors to share their stories, drawings, or reflections on cultural heritage sites through digital media platforms, exhibition panels, or physical walls. Similarly, the InCulture project [10] integrates intangible cultural heritage (ICH) into narratives and provides rich presentation formats, such as multimedia and AR, while encouraging user participation in content creation and community interaction. In museum experience design, the QRator project [2] introduces the concept of interactive labels, each displaying a QR code. Visitors can scan the codes with their smartphones to access information and discussions related to the exhibits. Interaction in this process is bidirectional, as visitors not only passively receive information but also contribute to the creation of exhibit interpretations. Instead of replacing physical artifacts, the use of digital technology in museums serves as an extension that enhances the visitor experience [4].

3 Method

Rooted in the practice of participatory ethnography, within the general framework of constructive design research and research-through-design methods, this article presents a design project as a case study where interactive digital storytelling is applied to create personal narratives for museum experiences [11–13]. The design of this project has been thoroughly documented and reflected upon throughout its development, it has been recorded, reflected on, and later became the data source for qualitative analysis.

4 Design Practice

MuseCode is a museum experience design that enables visitors to create personal narratives through interactive digital storytelling both on-site and off-site (see Fig. 1) in the context of the "cultural heritage boom" in China. This section will first illustrate MuseCode system and its components, and then detail the underlying mechanisms of each phase.

Fig. 1. The MuseCode system overview.

4.1 Components

The MuseCode system comprises three main components including a mobile App, an automated souvenir printer, and a souvenir package. The mobile App features three core functions: collecting cultural artifacts, creating a personal digital collection gallery, and scanning souvenir cards to access additional related knowledge (refer to Fig. 2). The automated printer operates based on advanced printing techniques, with the printing process primarily executed by robotic arms. As Fig. 3 shows, visitors can observe the entire production process through 360-degree glass panels, providing an opportunity to learn about and understand the intricacies of printing craftsmanship. Finally, visitors will receive a semi-finished souvenir of a specific artifact, a boxed set of four cards and a display stand. The four cards include the outer contour card, the inner contour card, the texture card, and the base color card.

Fig. 2. Example interfaces in the MuseCode App.

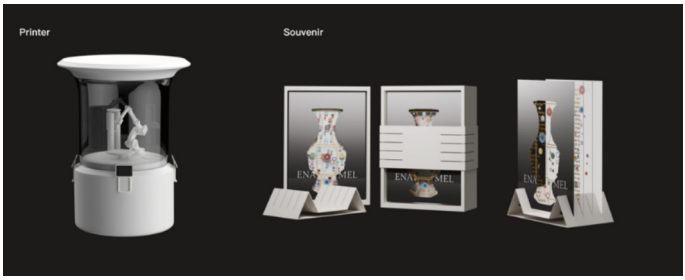


Fig. 3. The automated printer and the souvenir package.

4.2 MuseCode Workflow and Usage Scenario

The entire using process is displayed in Fig. 4. When visitors are on-site, they can use a mobile App to scan the ID number of the exhibits they are interested in and collect them. All artifacts visitors collect will form a personal electronic library, allowing them to access it anytime and anywhere. Then visitors can use the automated MuseCode printer to convert their digital collectibles into physical souvenirs. During the customization process, artificial intelligence (AI) is used to analyze and extract the characteristics of the selected exhibit into four layers: outer contour, inner contour, texture, and base color. As Fig. 4 shows, we take the enamel vase from the Palace Museum as an example to explain the conversion process. It begins by extracting the outer contour of the vase from a front-facing perspective. This contour is printed onto a card and further processed using laser engraving technology to create a hollowed-out silhouette card. Subsequently, other semantic features of the artifact are extracted layer by layer. Finally, they will receive a boxed set of four cards for them to assemble into a complete souvenir.

Additionally, this design extends visitor experiences to the off-site phase, encouraging visitors to further explore the cultural heritage. They can scan any of the four cards using the MuseCode App off-site to obtain more information from rich dimensions. For example, when scanning the texture card, they can delve into the history and craftsmanship of this texture and access the texture database, where they can view all artifacts unearthed so far with the same texture, excluding those they have already added to their collection. This interconnected presentation of related information can spark users' curiosity and inspire them to embark on the next plan of their museum journey.

Additionally, when scanning the fully assembled set of four cards, they can browse detailed information about this cultural relic and view its AR model.

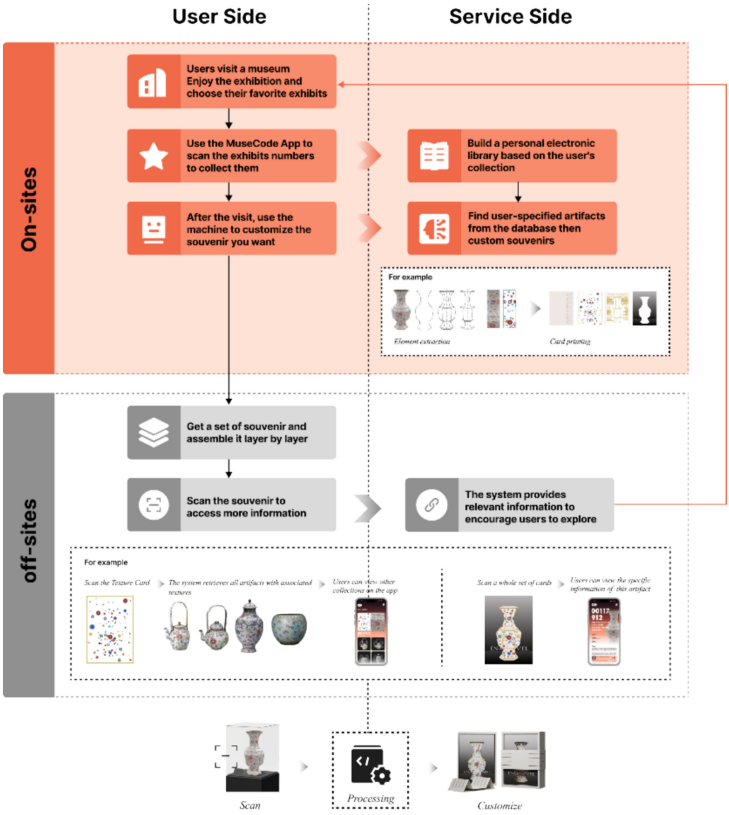


Fig. 4. Interactive using process chart of MuseCode.

5 Reflections

5.1 AI Deconstruction as a Creation Method

AI technology revolutionizes the transformation of cultural relics into souvenirs by deconstructing artifacts into four conceptual layers, each representing a distinct dimension—historical, aesthetic, functional, or symbolic. This approach moves beyond traditional 1:1 replication, instead focusing on abstraction and synthesis to create meaningful representations of cultural heritage. Each layer captures the essence of the artifact through descriptive forms, which are integrated into a multi-layered souvenir, offering visitors a personalized and emotional connection to heritage. This innovative method redefines the experience of cultural artifacts, prioritizing creativity and interaction over strict fidelity. The design philosophy emphasizes abstraction and reinterpretation, making it universally

adaptable to diverse cultural relics, including calligraphy, paintings, vessels, garments, and sculptures. MuseCode employs AI deconstruction to preserve the narrative depth of artifacts while enhancing their relevance and accessibility in contemporary cultural and creative design. By shifting the focus from authenticity to imaginative engagement, this method bridges traditional heritage with modern design practices, fostering a fresh appreciation for cultural relics.

5.2 Souvenir as a Gateway

Scholars [14, 15] emphasize the museum store's role in enhancing the overall museum experience, highlighting visitors' intentions to explore and purchase souvenirs. In the MuseCode project, souvenirs transcend their traditional role by serving as gateways to personal narratives. Visitors can scan souvenirs to access personalized libraries, blending tangible artifacts with intangible cultural heritage. The project links physical artifacts with layered heritage meanings, using material elements reproduced on four layers of selected antiques as aesthetic and narrative anchors. These layers evoke cultural, historical, and emotional connections, deepening engagement with the stories behind the objects. By merging physicality with virtual experiences, MuseCode underscores materiality's importance in digital heritage engagement. Designing souvenirs as interactive gateways ensures that even in virtual spaces, the sensory and emotional resonance of physical artifacts fosters meaningful connections to cultural heritage.

5.3 Personal Narrative as Cohesion

MuseCode leverages modern computing and interactive technologies to strengthen connections between people, physical spaces, and digital information, enriching visitors' cultural experiences. By integrating various aspects of the museum visit into a seamless narrative, it enables personalized storytelling through pre-visit, on-site, and post-visit interactions. Visitors engage actively by collecting artifacts, customizing souvenirs, and revisiting their collections, constructing unique personal narratives. Interactive digital storytelling fosters this shift from passive observation to active participation, making museum experiences more meaningful on an individual level. From a cultural communication perspective, this approach bridges gaps by creating an integrated system that connects every stage of the museum journey, offering a cohesive experience. In MuseCode's interactive narratives, audiences become active users, engaging with virtual exhibits through sensory operations, physical actions, and cognitive exchanges. This dynamic interaction ensures a holistic and effective realization of the museum narrative, blending technology with personal engagement.

6 Conclusion

In the digital era, museums have evolved from storytellers to story co-creators, collaborating with visitors to construct dynamic, participatory cultural experiences. This shift redefines museums as not just heritage preservers but interactive platforms blending physical and virtual spaces. Museums now emphasize intangible dimensions of cultural

heritage, such as linking generations, preserving collective memory, and fostering cultural identity. This research presents the MuseCode project as a case study and highlights interactive digital storytelling as a key method, summarized in three aspects including using AI deconstruction as a creative tool, designing souvenirs as gateways, and fostering personal narratives as cohesive elements. These insights inspire innovation in cultural and creative products while enhancing storytelling and personal connections to heritage.

References

1. Fu, K., Wu, R., Tang, Y., Chen, Y., Liu, B., Lc, R.: 'Being Eroded, Piece by Piece': Enhancing engagement and storytelling in cultural heritage dissemination by exhibiting GENAI Co-Creation artifacts. In: Designing Interactive Systems Conference (2024). https://doi.org/10.1145/3643834.3660711
2. Bailey-Ross, C., Gray, S., Ashby, J., Terras, M., Hudson-Smith, A., Warwick, C.: Engaging the museum space: mobilizing visitor engagement with digital content creation. Digit. Scholarsh. Humanit. fqw041 (2016). https://doi.org/10.1093/llc/fqw041
3. Trichopoulos, G., Aliprantis, J., Konstantakis, M., Michalakis, K., Caridakis, G.: Tangible and personalized DS application approach in cultural heritage: the CHATS project. Computers **11**(2), 19 (2022). https://doi.org/10.3390/computers11020019
4. Liu, P., Lan, L.: Museum as multisensorial site: story co-making and the affective interrelationship between museum visitors, heritage space, and digital storytelling. Mus. Manage. Curatorship **36**(4), 403–426 (2021). https://doi.org/10.1080/09647775.2021.1948905
5. Katifori, A., Karvounis, M., Kourtis, V., Perry, S., Roussou, M., Ioanidis, Y.: Applying interactive storytelling in cultural heritage: opportunities, challenges and lessons learned. In: Lecture Notes in Computer Science, vol. 603, pp. 603–612. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-030-04028-4_70
6. Bruner, J.: The narrative construction of reality. Crit. Inq. **18**(1), 1–21 (1991). https://doi.org/10.1086/448619
7. Alcantud-Díaz, M., Vayá, A.R., Gregori-Signes, C.: 'Share your experience': digital storytelling in English for tourism. Ibérica **27**, 185–204 (2014). http://www.aelfe.org/documents/09_27_Alcantud.pdf
8. Falco, F.D., Vassos, S.: Museum experience design: a modern storytelling methodology. Des. J. **20**(sup1), S3975–S3983 (2017). https://doi.org/10.1080/14606925.2017.1352900
9. Kim, S.: Virtual exhibitions and communication factors. Mus. Manage. Curatorship **33**(3), 243–260 (2018). https://doi.org/10.1080/09647775.2018.1466190
10. Mathioudakis, G., et al.: InCulture: a collaborative platform for intangible cultural heritage narratives. Heritage **5**(4), 2881–2903 (2022). https://doi.org/10.3390/heritage5040149
11. Frankel, L., Racine, M.: The Complex Field of Research: for Design, through Design, and about Design. DRS Digital Library. https://dl.designresearchsociety.org/drs-conference-papers/drs2010/researchpapers/43
12. Blomberg, J., Karasti, H.: Ethnography: Positioning ethnography within participatory design. In: Simonsen, J., Robertson, T. (eds.) Routledge International Handbook of Participatory Design, pp. 86–116. Routledge, New York (2012)
13. Koskinen, J., Zimmerman, T., Binder, J., Redstrom, J., Wensveen, S.: Design research through practice: from the lab, field, and showroom. IEEE Trans. Prof. Commun. **56**(3), 262–263 (2013). https://doi.org/10.1109/tpc.2013.2274109

14. Balermpas, A., Manola, M.: Souvenirs: their role in tourism industry, cultural heritage and their utilization in experiential teaching. Arch. Bus. Res. **8**(9), 1 (2020). https://doi.org/10.14738/abr.89.8970
15. Kent, T.: The role of the museum shop in extending the visitor experience. Int. J. Nonprofit Voluntary Sect. Mark. **15**(1), 67–77 (2009). https://doi.org/10.1002/nvsm.368

Check for updates

A Semi-Automatic Text Recognition Tool for Pre-Colonial Handwritten Manuscripts in Devanāgari Script

Bharath Valaboju, Shagun Dwivedi, Kartik Chincholikar, Kaushik Gopalan(✉), and Vinod Vidwans

FLAME University, Pune 412115, India
{bharath.valaboju,shagun.dwivedi,kartik.chincholikar,
kaushik.gopalan,vidwans}@flame.edu.in

Abstract. Manual annotation of undigitized manuscripts is a resource and labor intensive endeavour. On the other hand, using out-of-the-box Optical Character Recognition (OCR) tools to digitise text from handwritten manuscripts is challenging, particularly in languages like Sanskrit, where the scripts have evolved significantly over time.

This poster presents an annotation tool which allows the user to extract text from undigitized manuscripts using OCR, following which users can make corrections to the OCR-detected text. Users can then request fine tuning on a few pages corrected by them, making the annotation process easier and more efficient for the subsequent pages by improving OCR performance.

Additionally, an issue faced in the annotation of manuscripts is that the Devanāgari script is sometimes laborious to edit due to the behaviour of conjunct clusters and the halant, necessitating many keystrokes for simple edits. To mitigate this issue, the application uses an additional text box that represents the text in the Harvard-Kyoto transliteration scheme, which is easier to edit given that most keyboards use the Roman alphabet. The tool also supports rare Sanskrit characters, which are not usually supported by standard typing tools.

The frontend, which is designed to reduce effort on the annotator's part by placing the recognised text in editable text boxes right below the line images, is implemented in Vue.js. The backend is a Flask server employing PyTorch for inference and fine-tuning. The application aims to facilitate efficient annotation by having users upload manuscript images, which are then processed by the backend, which segments text lines from the leaves and recognises the text in the line images. Initial evaluations from beta testers suggest that the application significantly enhances the efficiency of the annotation process, making it more accessible and reducing the time required for its completion.

Keywords: Sanskrit Manuscripts · GUI · Handwritten text recognition · AI Assisted Transcription

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 152–160, 2025.
https://doi.org/10.1007/978-3-031-94171-9_13

1 Introduction

Transcribing historical manuscripts manually is a tedious and resource intensive task. Out-of-the-box OCR solutions aim to speed up the process but despite significant advances in techniques to digitise text from handwritten documents over the decades, handwritten manuscripts remain a difficult challenge due to irregular layouts, marked by footnotes and doodles [16] [18] [20] [24]. Fading ink, smudges, paper wear and other forms of degradation can add noise to the image and make character recognition harder [11]. The lack of training data specific to Sanskrit manuscripts further exacerbates this issue, making it harder to improve model performance.

The tool proposed in this poster enables users to overcome the above challenges to semi-automatically transcribe text from Sanskrit manuscripts through an intuitive interface, allowing them to automatically or manually segment line images from leaves, and then recognize text from the line images using a dedicated iteratively fine-tuned text recognition model for each target manuscript.

2 Related Work

Out-of-the-box OCR frameworks such as DocumentAI [15] have difficulty in digitizing text from historical handwritten manuscripts. However, specialized frameworks for the same have been developed, which broadly perform the task in two steps: (a) layout analysis and text-line segmentation and (b) recognizing the text content from the segmented lines images. Recent notable approaches for segmenting line images from historical Sanskrit palm leaf manuscripts include SeamFormer [26] and Palmira [23]. Palmira [23] also introduced the Indiscapes2 dataset, which offers a significantly more diverse collection of handwritten Indic manuscripts compared to its predecessor. Jindal and Ghosh [17] were the first to employ a faster region-based convolutional neural network (R-CNN) method for robustly segmenting text lines in an ancient handwritten document written in Devanagari script. For recognizing text content from line images, EasyOCR [5] provides support for 80+ languages in printed form of their respective scripts, with support for Handwritten Text Recognition still in the pipeline. Their OCR models can be easily fine tuned with a small amount of segmented line images to recognise handwritten text with great accuracy.

Transkribus [8] is a handwriting recognition software that can be trained to recognise both handwritten historical documents and printed texts, however it is only available as a paid service. eScriptorium [19], an open source, is a web-based software designed for automated or manual segmentation and text recognition for Arabic, Hebrew, Greek and Latin. It provides for manual correction of layout analysis and text segmentation, which can be used to train a text detection model. For this, the platform uses the open-source HTR engine, Kraken. Handwriting recognition and post-correction frameworks tailored for Devanagari script have been developed by Saluja et al., [21,22], Adiga et al., [10], Das [14], and Vinitha and Jawahar [27]. These frameworks assist human

annotators in digitizing text by dynamically generating domain-specific dictionaries, clustering similar recurring errors, and automatically identifying misspelled words. OpenOCRCorrect [22] is an end-to-end Indic OCR framework for error detection and corrections by providing suggestions for words incorrectly recognised by the OCR model to the user. The framework has been built under the umbrella project Akshar Anveshini [1], which aims to make historic books easily accessible and discoverable to a greater number of people than a physical print format allows for. ind.senz [6] is an OCR program for Sanskrit, that converts printed Sanskrit texts into computer readable, editable and searchable digital documents in Unicode-Devanāgari encoding; however, it is also only available as a paid service. HInDoLA [25], or Historical Intelligent Document Layout Analytics, is a web-based system consisting of an Annotation Tool, Dashboard Analytics and Machine Learning engines, for handwritten historic documents.

In this paper, we propose an open-source, web-based Sanskrit manuscript annotation tool with support for both automatic and manual line segmentation, and text recognition from segmented line images. It provides a user-friendly experience for the user when making corrections to the recognized text, and also supports iterative fine tuning of the text recognition model on target manuscripts.

3 Methodology

OCR reliability can be affected by factors such as variations in handwriting, long term deterioration, and irregular layouts. This tool addresses these challenges by allowing users to correct a subset of the recognized text, which in turn enhances the accuracy of inferred text across the entire document. The tool offers a pre-trained model from Jaided AI [4] and, if previously used, fine-tuned models trained by other users. Users can upload their material as JPG files and select a model to run OCR. The tool then runs a segmentation model powered by CRAFT and segments the images into lines. Following this, the selected recognition model is run on these images and the inferences are displayed to the user. The user can make corrections to a subset of the recognized text and then request a finetune based on their correction. The tool fine-tunes the selected model using the user's corrections as ground truth. The user can then use this finetuned model to get more accurate inferences on the manuscript it is fine-tuned on. The user flow can be seen in Fig. 1. This approach significantly reduces manual effort and improves efficiency. The tool is a RESTful web application built in Vue.js and Flask. The underlying segmentation and recognition models are powered by PyTorch. The technicalities will be detailed in the following sections. The application's code can be accessed at its GitHub repository.

3.1 Frontend

The application's frontend is a web application written in Vue.js 3 [9]. No additional download apart from a web browser is required for the frontend. Upon opening the application, the user is offered the choice between uploading a new

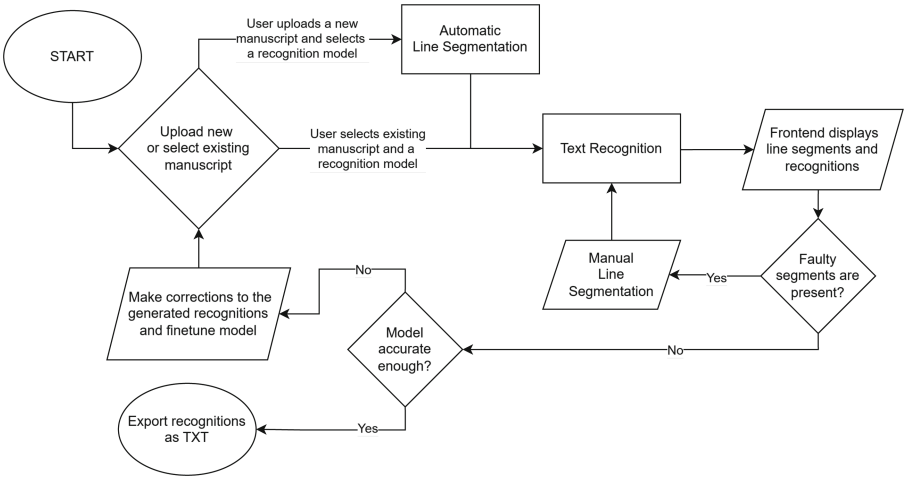


Fig. 1. User flow diagram.

manuscript or running inference on a previously uploaded one. In case the user opts to upload a new manuscript, they are prompted to upload images of the manuscript in JPG format and to select the pre-trained model to use for inference. Otherwise, they can select a previously uploaded manuscript along with a model, which is ideally the model fine-tuned on that manuscript to use for inference. The application then recognises the text content for each line image for each page of the manuscript as shown in Fig. 2.

In case the application doesn't segment the line images properly, the user can segment the line images manually as shown in Fig. 4. Once the line images are re-segmented then the recognition model can be run again, to recognize the text content from the line images. The user can then make corrections to a subset of the recognized text and then request a fine-tuning, which will improve its text recognition performancce on the subsequent pages. There is also a provision to toggle a secondary textbox that contains a representation of the primary textbox in the Harvard-Kyoto scheme and is synchronized with it. This enables convenient editing and circumvents an issue that involves conjunct clusters incorrectly sticking to each other when backspacing. The user can also export the inferences to TXT format. Additionally, the interface has a palette that lets users insert rare Devanāgari characters. Both of these features are shown in Figs. 2 and 3 respectively.

3.2 Backend

The tool's tech stack consists of a REST API architecture powered by the Python web micro-framework Flask. Flask has been chosen as it is a Python framework, which enables working with its machine learning ecosystem. Flask has the added advantage of enabling rapid development while also allowing for finer control over

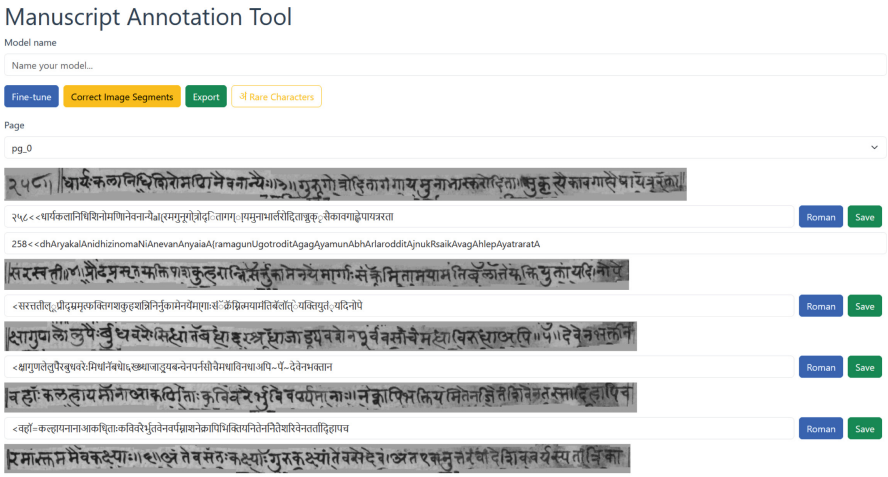


Fig. 2. Here, users can make corrections to the recognitions made by the tool and request to finetune the recognition model. There is a toggle for Harvard-Kyoto input as shown in the first text box in the image. Users can also export the recognitions to TXT in this page.

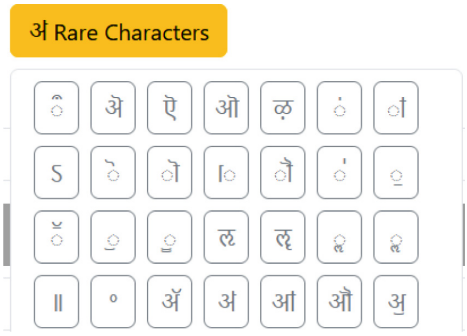
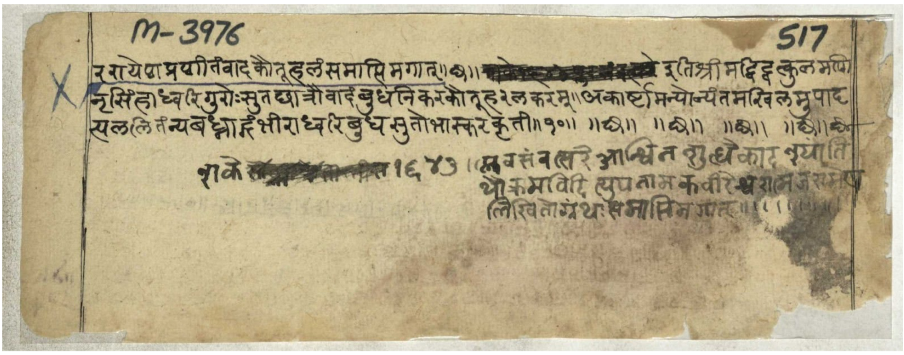


Fig. 3. This palette can be used to copy rare characters to clipboard.

the specifics. SQLite [7] is being used for the database owing to its simplicity. The application makes use of the REST API architecture. It has been chosen as it is the industry standard and makes for a reasonably straightforward development experience. Text recognition in the manuscripts is done in two steps: The first step consists of segmentation of lines, and the second step consists of recognition of the text content of the lines. The segmentation method has Character-Region Awareness For Text detection(CRAFT) [12] as the backbone, whereas the recognition model is powered by a CNN-BiLSTM-CTC neural network architecture implemented by Clova AI [3].

Generalised robust text line segmentation remains challenging when dealing with diverse page layouts, as shown by recent studies [16] [18] [20] [24].



(a) An edge case leaf with an unexpected layout where automatic line segmentation fails, necessitating manual line segmentation as shown in (b).

(b) User interface for manually identifying and segmenting lines from a leaf, where all the characters in a line are marked with the same colour.

Fig. 4. Manual Line Image Segmentation: (a) an edge case leaf where the automatic line segmentation fails, and (b) manually segmenting the lines from the same leaf.

To address this challenge, we limit the automatic line segmentation capabilities of this tool to manuscripts which meet a layout-based selection criterion: manuscripts which contain no pictorial elements and are single-columned texts with no line breaks. While we observe that a vast majority of handwritten pre-colonial Sanskrit manuscripts which we have encountered fit this selection criteria, we often face edge case pages which do not meet these criteria. For such pages, we enable the user to manually annotate the lines from the page as seen in Fig. 4.

When automatically segmenting line images from page images, we propose an approach that builds on the scene text detection method CRAFT [13]. CRAFT converts an image to a heatmap, where the regions with characters present have a high temperature. Next, we use a piecewise projection profile based approach

to group all characters of a line together and then segment the line images. This approach can successfully handle irregular curved handwritten lines, as projection profiles are taken at the left, centre and right regions of the page image. We noticed that applying projection profile based approaches on the heat maps is significantly easier compared to the original images, as the heat map does not contain pictures, smudges, ink blots, page texture, page tears and other non-text elements.

When manually segmenting line images, the CRAFT heatmap is converted to a set of 2D points, where each point denotes the presence of a character. The user then manually hovers the mouse over all the points contained in an individual line and marks them with a common colour, as seen in Fig. 4b. Next, these annotations and the heatmap are used to extract line images from the original image.

Once the line images are segmented, the recognition model is used to recognise the text content from the segmented line images. The base recognition model used is a pre-trained Devanāgari CNN-BiLSTM-CTC model trained by EasyOCR [4]. We observe that the recognition accuracy increases significantly when the base model is fine-tuned on as few as one or two pages of the target manuscript. In other words, the input-label pairs (segmented line images, ground truth annotations) obtained through user OCR-corrections are used to fine-tune the pre-trained recognition model for the Devanāgari script. Thus, we maintain a dedicated recognition model for each manuscript, which is periodically fine-tuned as more pages get transcribed.

4 Discussion and Future Work

With this tool, we enable the digitisation of handwritten manuscripts using a two-step process. The first step consists of segmentation of lines, and the second step consists of recognition of the text content of the lines. We also enable the recognition model to be fine-tuned on corrected ground truth transcriptions - allowing it to improve its recognition performance further.

In the future, the tool can benefit from a third step, in which the language model ByT5-Sanskrit, along with existing post-correction methods, can be used to perform the task of making automatic corrections to the transcribed text. The tool currently only supports segmenting lines from page images, without the ability to label regions of the pages, such as images, columns, text blocks, page numbers, marginal notes, etc. Hence, it might be beneficial in the future to adhere to the standardization enabled by Segmonto [2], which provides a controlled vocabulary that describes the layout of a page. Adding support for deep annotation features such as named entity recognition and keyword spotting could also be beneficial.

Acknowledgments. The authors wish to express their gratitude to Ansh Kushwaha and Rishav Jain for their support and their work on which this tool builds upon. The authors are thankful for the valuable user insights and annotations provided by Dr.

Tarinee Awasthi. The authors also wish to express their thanks and appreciation to Lalchand Research Library, DAV College, Chandigarh, India, for uploading the pre-colonial Sanskrit palm leaf manuscript images online and making them available for educational and research purposes.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Akshar anveshini homepage. https://www.cse.iitb.ac.in/~ocr/
2. ALTO - SegMonto documentation. https://segmonto.github.io/pp/alto/
3. Deep text recognition benchmark. https://github.com/clovaai/deep-text-recognition-benchmark
4. EasyOCR devanāgarī model. https://github.com/JaidedAI/EasyOCR/releases/download/pre-v1.1.6/devanagari.zip
5. EasyOCR repository. https://github.com/JaidedAI/EasyOCR
6. OCR and digitization software for Hindi and Sanskrit - ind.senz. https://www.indsenz.com/int/index.php
7. SQLite home page. https://sqlite.org/
8. Transkribus homepage. https://www.transkribus.org/
9. Vue.js - the progressive JavaScript framework. https://vuejs.org/
10. Adiga, D., et al.: Improving the learnability of classifiers for Sanskrit OCR corrections. In: The 17th World Sanskrit Conference, Vancouver, Canada. IASS, pp. 143–161 (2018)
11. AlKendi, W., Gechter, F., Heyberger, L., Guyeux, C.: Advancements and challenges in handwritten text recognition: a comprehensive survey. J. Imaging **10**(1), 18 (2024). https://doi.org/10.3390/jimaging10010018
12. Baek, J., et al.: What is wrong with scene text recognition model comparisons? Dataset and model analysis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4715–4723 (2019). https://doi.org/10.1109/ICCV.2019.00481
13. Baek, Y., Lee, B., Han, D., Yun, S., Lee, H.: Character region awareness for text detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9365–9374 (2019). https://doi.org/10.1109/CVPR.2019.00959
14. Das, D.: Enhancing OCR Performance with Low Supervision. Ph.D. thesis, International Institute of Information Technology Hyderabad (2021)
15. Hegghammer, T.: OCR with Tesseract, Amazon Textract, and Google document AI: a benchmarking experiment. J. Comput. Soc. Sci. **5**(1), 861–882 (2022). https://doi.org/10.1007/s42001-021-00149-1
16. Janes, J., Pinche, A., Jahan, C., Gabay, S.: Towards automatic TEI encoding via layout analysis. In: Fantastic future 21, 3rd International Conference on Artificial Intelligence for Librairies, Archives and Museums (2021)
17. Jindal, A., Ghosh, R.: Text line segmentation in Indian ancient handwritten documents using faster R-CNN. Multimedia Tools Appl. **82**(7), 10703–10722 (2023). https://doi.org/10.1007/s11042-022-13709-y
18. Kiessling, B., Kurin, G., Miller, M.T., Smail, K.: Advances and limitations in open source Arabic-script OCR: a case study. arXiv preprint arXiv:2402.10943 (2024). https://doi.org/10.48550/arXiv.2402.10943

19. Kiessling, B., Tissot, R., Stokes, P., Stokl Ben Ezra, D.: eScriptorium: an open source platform for historical document analysis. In: 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), pp. 19–19. IEEE, Sydney, Australia (2019). https://doi.org/10.1109/ICDARW.2019.10032, https://ieeexplore.ieee.org/document/8893029/
20. Nikolaidou, K., Seuret, M., Mokayed, H., Liwicki, M.: A survey of historical document image datasets. Int. J. Doc. Anal. Recogn. (IJDAR) **25**(4), 305–338 (2022). https://doi.org/10.1007/s10032-022-00405-8
21. Saluja, R., Adiga, D., Chaudhuri, P., Ramakrishnan, G., Carman, M.: Error detection and corrections in Indic OCR using LSTMs. In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), vol. 1, pp. 17–22. IEEE (2017). https://doi.org/10.1109/ICDAR.2017.13
22. Saluja, R., Adiga, D., Ramakrishnan, G., Chaudhuri, P., Carman, M.: A framework for document specific error detection and corrections in Indic OCR. In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), vol. 4, pp. 25–30. IEEE (2017). https://doi.org/10.1109/ICDAR.2017.308
23. Sharan, S.P., Aitha, S., Kumar, A., Trivedi, A., Augustine, A., Sarvadevabhatla, R.K.: Palmira: a deep deformable network for instance segmentation of dense and uneven layouts in handwritten manuscripts. In: Lladós, J., Lopresti, D., Uchida, S. (eds.) ICDAR 2021. LNCS, vol. 12822, pp. 477–491. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86331-9_31
24. Simistira, F., Seuret, M., Eichenberger, N., Garz, A., Liwicki, M., Ingold, R.: DIVA-HisDB: a precisely annotated large dataset of challenging medieval manuscripts. In: 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 471–476. IEEE (2016). https://doi.org/10.1109/ICFHR.2016.0093
25. Trivedi, A., Sarvadevabhatla, R.K.: HInDoLA: a unified cloud-based platform for annotation, visualization and machine learning-based layout analysis of historical manuscripts. In: 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), vol. 2, pp. 31–35. IEEE (2019).https://doi.org/10.1109/ICDARW.2019.10035
26. Vadlamudi, N., Krishna, R., Sarvadevabhatla, R.K.: SeamFormer: high precision text line segmentation for handwritten documents. In: Fink, G.A., Jain, R., Kise, K., Zanibbi, R. (eds) International Conference on Document Analysis and Recognition, pp. 313–331. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-41685-9_20
27. Vinitha, V., Jawahar, C.: Error detection in Indic OCRs. In: 2016 12th IAPR Workshop on Document Analysis Systems (DAS), pp. 180–185. IEEE (2016). https://doi.org/10.1109/DAS.2016.31

Innovative Dissemination Path of Yi Character in Human-Computer Interaction: A Case Study of Personalized Cultural Product Customization

Ziyang Wang and Yuanfang Zhao(✉)

College of Art and Design, Shenzhen University, Shenzhen 518060, People's Republic of China
zhaoyuanfang@szu.edu.cn

Abstract. The Yi people are one of the ethnic minorities in China. The Yi character, with its unique character structure, rich historical connotation, and profound ethnic cultural heritage, holds significant value for academic research and cultural dissemination. However, with the popularization of simplified Chinese characters, the number of people capable of recognizing and understanding Yi character has been decreasing, and this precious cultural heritage is quietly fading from the collective memory of the Yi people. This study aims to explore innovative strategies for the graphical design of Yi character, transforming traditional Yi characters into graphic symbols with modern aesthetics through contemporary design language, and subsequently developing a series of cultural creative products. Through the integration of culture and design, the Yi character is endowed with new visual expressions while promoting its application and dissemination within a multicultural context.

Using the KANO model to analyze the needs of visitors and consumers regarding Yi character cultural products, this study derives the following key conclusions: (1) The KANO model effectively identifies different user needs for Yi character cultural products. Consumers developed an emotional connection to Yi culture during personalized customization, providing an important basis for the precise positioning of cultural creative products. (2) The introduction of customization options significantly enhances user satisfaction with cultural products. Users particularly appreciated personalized involvement in adjustable features like size and color, effectively increasing the Product Index (PI) of cultural products. (3) Through interactive platform design, users could not only participate in the product design process but also deepen their understanding of the meanings of Yi character during the interaction, thereby reducing cross-cultural cognitive barriers to some extent and promoting cross-cultural communication. Cultural products can establish a positive interaction between commercial value and cultural preservation, promoting the sustainable development of intangible cultural heritage. This study provides a practical reference for the dissemination and innovative application of Yi character in the contemporary design context, aiming to demonstrate its application value in the interaction design of personalized customization and cultural dissemination.

Keywords: Yi Character · Graphical Design · Cultural Creative Products · Cross-Cultural Design · Personalized Customization

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 161–171, 2025.
https://doi.org/10.1007/978-3-031-94171-9_14

1 Introduction

The Yi people in China are mainly distributed in Yunnan, Sichuan, Guizhou, and Guangxi provinces. Among these regions, Liangshan Yi Autonomous Prefecture is the largest inhabited area of the Yi people and the key region for Yi ancient books and documents [1]. After the founding of the People's Republic of China—with the standardization of ethnic names and the implementation of national policies on nationality, language, and writing—the inherent characters of the Yi people were collectively referred to as "Yi characters" (or simply Yi characters). These characters were recorded on various media (such as bone, stone, gold, cloth, silk, bamboo, wood, leather, and paper), forming an extensive corpus of ancient Yi literature [2]. As an indigenous writing system, the Yi language is one of the oldest among China's national scripts and an important component of the nation's cultural diversity; it plays a unique role in studying the evolution of human civilization and in demonstrating the world's cultural diversity [3]. Although Yi characters are similar to Chinese characters in being mainly ideographic, they possess distinctive features. Some experts even regard ancient Yi characters as one of "the six ancient Chinese characters in the world," along with Oracle Bone Inscriptions, Sumerian, Egyptian, Mayan, and Halaba [4]. Liangshan Yi people call their language "Nuosu Buma," a term that conveys graphic as well as hieroglyphic meaning along with high aesthetic and artistic value [5]. The beauty of Yi characters is reflected in their design—the simple lines and harmonious compositions imbue them with ethnic characteristics and an abstract artistic quality. The unique symbolic nature of the Yi language enables designers to transform these characters into graphical elements for creative use.

Through a field visit to Liangshan Yi Autonomous Region in Sichuan Province, the author observed that the Yi language enjoys widespread public application locally (see Fig. **1**). Whether on shop facades, in slogans of various government agencies, or on everyday clothing, Yi is the predominant form of writing. This phenomenon reflects both its high prevalence in daily life and its significant role in expressing local cultural identity and regional characteristics. As a symbol of regional culture, the Yi language not only carries the mission of cultural inheritance but also exerts a subtle influence on the cultural belonging of local residents.

As early as 1979, Sichuan People's Broadcasting Station began broadcasting in Yi. Since 1976, the Liangshan National Film Translation Center has translated more than 1,000 films of various genres. In addition, Liangshan War TV Station not only launched a Yi language channel but also completed the translation of 110 TV dramas, movies, and TV films, thereby enriching the cultural life of the Yi people [6].

In the late 1990s, with the rapid development of network communication in China, the country truly entered the network era. However, information processing, speech recognition, and translation of the Yi language faced significant challenges. The computer-based Yi language information processing system has since been successfully developed, with systems such as the YWPS Yi language system, YWZN2000 Yi language system, and YW Yi language subtitle system emerging sequentially [7]. These technological breakthroughs not only promoted the spread and application of the Yi language in the information age but also laid an important foundation for the protection and innovation of Yi culture. With further globalization and digitalization, exploring the cultural value

Fig. 1. The Author Took a Real Shot in Butuo County, Liangshan Yi Autonomous Prefecture.

of the Yi language and its application in emerging fields supported by technology has become a new research direction.

2 Yi Graphic Cultural and Creative Products Demand Analysis

2.1 Sample Survey

The basic information of 206 valid samples was analyzed in terms of gender, age, and education level, as shown in Table 1. In general, the sample group is relatively young and highly educated, with enterprise employees and non-Yi individuals predominating.

2.2 KANO Analysis

KANO analysis is a model for understanding and analyzing customer needs, proposed by Japanese professor Noriaki Kano in 1984. It helps identify and classify different types of requirements, thereby facilitating more effective product or service design and improvement. The model divides user needs into five categories: essential demand, expected demand, attractive demand, indifference demand, and reverse demand [8].

Based on the pros and cons of the research questionnaire, the KANO model was applied to analyze consumers' attitudes toward each requirement under conditions of both provision and non-provision. The classification results are presented in Table 2.

According to the calculation method in Table 2, 14 design requirements for cultural and creative products were evaluated. The final classification results are shown in Table 3. Among these 14 requirements, 4 were categorized as essential, 3 as expected, 4 as attractive, and 3 as indifference requirements.

Table 4 is a summary of the design requirements of Cultural Creative products, including four essential requirements: style, material, color and technology; Three

Table 1. Basic Information of Respondents.

Characteristic	Group	Frequency	Percentage
Gender	Man	111	53.9
	Woman	95	46.1
Age	Under 20 Years Old	9	4.4
	20–29 Years Old	50	24.3
	30–39 Years Old	78	37.9
	40–49 Years Old	56	27.2
	50 Years Old And Above	13	6.3
Degree of Education	High School (Including Technical Secondary School) And Below	51	24.8
	College/Undergraduate Course	142	68.9
	Master Degree Or Above	13	6.3
Nation	Yi Ethnic Group	36	17.5
	Other	170	82.5
Occupation	Student	18	8.7
	Public Officials (Civil Servants/Institutions, Etc.)	29	14.1
	Employees/Staff of An Enterprise	113	54.9
	Individual Operator	31	15
	Other	15	7.3

expected needs: practicality, environmental protection and cultural characteristics; Four charm needs: attraction, uniqueness, inheritance and regionality; There are three indifferent needs: commemorative, emotional and educational.

2.3 Better-Worse Analysis

Product improvement processes may overlook the varying levels of user needs, thus missing opportunities to enhance overall satisfaction. To compensate for this, the study adopts the Better-Worse satisfaction coefficient - proposed by Berger C. et al. - based on the KANO model. This method refines the classification of demands and quantifies user satisfaction, thereby optimizing the analysis results.

If a given feature is provided, the improved satisfaction coefficient is calculated as:

$$\textbf{Better} = (\mathbf{A} + \mathbf{O})/(\mathbf{A} + \mathbf{O} + \mathbf{M} + \mathbf{I}).$$

If the feature is not provided, the reduced dissatisfaction coefficient is calculated as:

$$\textbf{Worse} = -(\mathbf{M} + \mathbf{O})/(\mathbf{A} + \mathbf{O} + \mathbf{M} + \mathbf{I}).$$

Table 2. Demand Calculation Method.

Provide	Does not provide					Demand category
	Be satisfied	Take…for granted	It doesn't matter	Tolerable	Dissatisfied	
Be satisfied	Q	A	A	A	O	M: essential requirements O: expected demand A: charm demand I: no difference in demand R: reverse demand Q: wrong result
Take…for granted	R	I	I	I	M	
It doesn't matter	R	I	I	I	M	
Acceptable	R	I	I	I	M	
Dissatisfied	R	R	R	R	Q	

Table 3. KANO Results.

Serial Number	Project	Q	R	A	I	O	M	Kano Demand
Q1	Style	1	9	29	58	29	80	M
Q2	Texture	3	23	38	46	23	73	M
Q3	Colour	4	11	41	44	26	80	M
Q4	Craft	1	19	36	49	27	74	M
Q5	Attraction	2	16	70	36	28	54	A
Q6	Practicability	2	20	36	28	78	42	O
Q7	Uniqueness	5	14	72	49	23	43	A
Q8	Environmental Protection	1	21	42	41	70	31	O
Q9	Culture Trait	1	16	40	55	71	23	O
Q10	Inheritance	1	16	72	53	30	34	A
Q11	Territoriality	6	14	66	47	42	31	A
Q12	Commemorative	3	15	34	96	21	37	I
Q13	Emotionality	3	22	33	83	17	48	I
Q14	Educational	2	14	39	82	37	32	I

The Better coefficient ranges between 0 and 1, with values closer to 1 indicating a greater impact on increasing user satisfaction, whereas values near 0 indicate a lesser impact. Similarly, the Worse coefficient ranges between –1 and 0; values closer to –1

Table 4. Statistics of Design Requirements Results.

Demand Type	Project	
M: essential requirements	Q1	Style
	Q2	Texture of Wood
	Q3	Colour
	Q4	Craft
O: expected demand	Q6	Practicability
	Q8	Environmental Protection
	Q9	Culture Trait
A: charm demand	Q5	Attraction
	Q7	Uniqueness
	Q10	Inheritance
	Q11	Territoriality
I: no difference in demand	Q12	Commemorative
	Q13	Emotionality
	Q14	educational

imply a greater effect on reducing dissatisfaction, while values near 0 indicate a smaller effect.

The specific Better-Worse coefficients of this study are shown in Table 5. As can be seen from Table 5, the Better coefficient indicates that if these needs can be met, consumer satisfaction will increase by the corresponding coefficient percentage. The Worse coefficient indicates that if these needs are not met, consumer satisfaction will decrease by the corresponding coefficient percentage.

3 Method

Yi Graphic Cultural and Creative Products Demand Analysis. Among the many rich cultural symbols, the author chose the five groups of words carrying the unique cultural significance and life philosophy of the Yi people. In order to make these words not only words, but visual symbols that can tell stories, designers need to conceive how to graph them in the design process, so that they can not only retain the original meaning, but also link with modern aesthetic interest.

Design Analysis of ‘ꆏꌠꁱꂷ’ (Bright Future). Through meticulous adjustments of the strokes in Yi script, the fluidity of curves and circles is enhanced, optimizing the font structure to achieve a more recognizable proportional relationship. By employing rhythmic line language and vibrant color schemes, the design constructs the imagery of vastness and brightness implied by “future.” Visual patterns incorporate elements of azaleas, symbolizing hope and prosperity (see Fig. 2), which not only highlight the

Table 5. Better-Worse coefficient.

Serial Number	Project	Kano Demand	Better	Worse	Better-Worse Demand
Q1	Style	M	0.296	−0.556	M
Q2	Texture of Wood	M	0.339	−0.533	M
Q3	Colour	M	0.351	−0.555	M
Q4	Craft	M	0.339	−0.543	M
Q5	Attraction	A	0.521	−0.436	A
Q6	Practicability	O	0.620	−0.652	O
Q7	Uniqueness	A	0.508	−0.353	A
Q8	Environmental Protection	O	0.609	−0.549	O
Q9	Culture Trait	O	0.587	−0.497	O
Q10	Inheritance	A	0.540	−0.339	A
Q11	Territoriality	A	0.581	−0.392	A
Q12	Commemorative	I	0.293	−0.309	I
Q13	Emotionality	I	0.276	−0.359	I
Q14	Educational	I	0.400	−0.363	I

Yi people's aspirations for a better life but also align with modern audiences' aesthetic demands for visual impact.

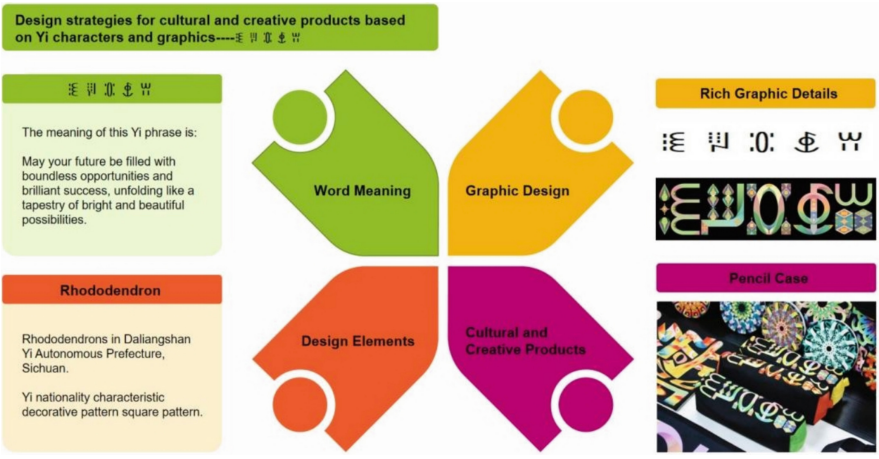


Fig. 2. A (Bright Future) Design Analysis.

Design Analysis of ‘ ꆏꌠꁱꂿ’ (Melodious). As a traditional musical instrument of the Yi people, the Kouxianqin bears hundreds of years of historical accumulation, is the cultural heritage of the ancestors of the Yi people, and contains rich cultural connotations. In many areas of Liangshan Yi Autonomous Prefecture, Kouxianqin is used as a tool to express feelings. In order to fit the musical instrument characteristics of the oral stringed instrument, the author has incorporated a large number of note elements into the “sweet” text design (see Fig. 3), combining the abstract strings with the gyrating sound wave form, so that the audience can perceive the fluidity and harmonious beauty of music.

Fig. 3. A ꆏꌠꁱꂿ(Melodious) Design Analysis

Design Analysis of ‘ ꂷꊰ’ (Peace). In graphic semiotics, the square represents stability, firmness and balance, which is one of the common basic geometric forms in nature and human civilization. It symbolizes the integrity and harmony of all directions, implying the stability of the earth and the order of the universe. The design of “Ping ‘an” tends to be square (see Fig. 4), which not only conforms to the preference of modern design for concise and lively forms, but also skilfully borrows the symbolic meaning of the square to convey the tranquillity, stability and shelter represented by “Ping ‘an”, which is highly consistent with the internal meaning of the word “Ping ‘an”. In essence, this design technique is to use visual language to strengthen the meaning function of words, so that even viewers who don’t understand Yi language can intuitively feel the meaning of security and stability. In Yi culture, the concept of square has a special position. The “four realms” or “land of four realms” in the Yi people’s cosmology are often closely related to the order of all things, the cycle of four seasons and the rhythm of life. “Eight Directions” is a traditional orientation sign, also known as eight directions, which helps Yi ancestors identify the direction and time.

Fig. 4. A (Peace) Design Analysis.

4 Conclusion

This study takes the graphic design of Yi characters as the core, combines the concepts of human-computer interaction and customized design, and explores the innovative path of combining culture and creative design. By designing and launching a series of cultural and creative products (see Fig. 5), these achievements were exhibited and sold through the T-fashion enterprise platform, which also attracted the attention and recognition of many consumers.

Fig. 5. Graphic Cultural and Creative Products of Yi Characters.

This series of cultural and creative products have been reported on Shenzhen Metropolitan Channel and Shenzhen Radio's Pioneer 898 (see Fig. 6). This research and practice verify the commercial potential of local characteristic culture in modern

commodity market, and also provide a feasible model for cultural inheritance and innovation. Future research can further explore the possibility of the integration of culture and technology, and open up more ways for the protection and development of multiculturalism.

Fig. 6. Images Broadcast by Metropolitan Channel and Shenzhen Radio Pioneer 898.

5 Conclusion

As the core cultural capital, Yi characters can be reproduced and spread in contemporary society through innovative interpretation of modern design language, and can also provide new opportunities for the development of cultural tourism industry in poor areas. As the precious wealth of the Chinese nation, local culture contains resources such as geographical landscape, historical culture and folk customs. Realizing the visual application of Yi characters in cultural and creative products can help the economic development of poor areas and promote the coordinated progress of urban and rural communities under the background of cultural tourism, cultural education and the promotion of local characteristic industries. This diversified application embodies the leap of Yi language from tradition to modernity, and lays a solid foundation for its extensive application in digital and graphic design and other emerging fields. This also provides valuable experience for exploring how minority languages can be more widely integrated into cultural and creative industries.

Future research can further focus on the deep combination of culture and technology, such as expanding the possibility of cultural expression through cutting-edge technologies such as artificial intelligence and virtual reality. At the same time, the user experience and market feedback can be analysed more systematically, and the design process of cultural and creative products can be optimized to meet the needs of different cultural backgrounds and consumer groups. Through more targeted research and design, we can not only deepen the interpretation and innovation of traditional culture, but also enhance the international influence of local culture under the background of globalization and provide more powerful support for cultural protection and multicultural development.

Through the exhibition and market application of cultural and creative products with graphic design of Yi characters, this study preliminarily verified the certain relationship

between cultural innovation and sustainable development of urban and rural areas. The promotion of cultural and creative products not only provides a new way for the inheritance of Yi culture, but also promotes the development of rural economy in the follow-up research and promotion to a certain extent. By combining traditional culture with modern design, Cultural Creative Products not only promote consumers' understanding and interest in Yi culture, but also inject vitality into the economy of poor areas and enhance local cultural self-confidence and economic sustainable development ability. The combination of cultural innovation and local characteristic resources shows the symbiotic relationship between urban and rural areas and provides new ideas for cultural tourism and local economic development.

Acknowledgements. The authors acknowledge the funding support by Shenzhen University Graduate Innovation Cultivation Program (Project No: 000066010688), Youth Fund of Humanities and Social Sciences Research of the Ministry of Education (Project No: 24YJCZH449), Guangdong Province Philosophy and Social Science Planning Project (Project Number: GD21YYS06).

References

1. Lu, A.X.T.: Historical value, protection status and system design of Yi literature. Univ. Libr. Work **38**(05), 50–53 (2018)
2. Fang, G.: Historical Draft of Yi Nationality, p. 5. Sichuan Nationalities Publishing House, Chengdu (1984)
3. Lu, Y., Li, W.: A summary of the compilation of Yi language dictionaries in China. J. Guizhou Inst. Eng. Appl. Technol. **37**(02), 99–100 (2019)
4. Lu, S.: Yi Ancient Books Are the Key Factor in Studying the Yangtze River Civilization and the History of Ethnic Groups in Southwest China (2017). http://mp.weixin.qq.com/s/q1gzFGHNiighdbi_5a11kg
5. Long, F.: Research on innovative design of Yi characters. Art Technol. **28**(12), 88 (2015)
6. A Color of A Min, Yang, Q.: A brief analysis of the process and development bottleneck of Yi language information dissemination in Liangshan Yi area. Urban Tour. Res. **07**(02), 37–42 (2012)
7. Ma, N.: Research on the application of visual graphic design of Yunnan minority characters. Hunan Packag. **06**(34), 132–136 (2021)
8. Mahai, R.-G., Aniu, M.-Z.: On the historical origin and cultural value of Yi peoples' four-square and eight-digit. J. Xichang College Soc. Sci. Edn. **04**(28), 6–13 (2016)

Application of K-means Algorithm in Extracting Thangka Pattern Colors for Cultural and Creative Product Design

Bin Wang[1,2], FangChao Yang[1,2(✉)], Jian Feng Wang[1,2], and Na Liang[1,2]

[1] Hanyang University, Ansan 15588, Republic of Korea
wangbin19045@hanyang.ac.kr, 2131763@tongji.edu.cn
[2] College of Design and Innovation, Tongji University, Shanghai, China

Abstract. Thangka, also referred to as Tangka, is a traditional scroll painting typically rendered on specially prepared fabrics or paper, primarily depicting the historical and cultural narratives of Buddhism. As one of the first groups of National Intangible Cultural Heritage in China, Thangka faces challenges in transmission within contemporary society. These difficulties arise from the extended duration of master-apprentice training, the significant reliance on individual artistic skill and emotional expression, and the declining interest among younger generations—factors that collectively hinder large-scale dissemination. To address these challenges, this study proposes employing scientific algorithms to systematically extract color and pattern elements from Thangka images and integrate them into modern cultural and creative product design. The research utilizes the K-means clustering algorithm to extract key color and pattern information from representative Thangka images and applies these data to the design and development of new products, thereby balancing traditional artistic value with contemporary market demands. The scope of this study includes color extraction, color analysis, and the exploration of these elements within modern design practices. By combining digital methods with artistic innovation, this study aims to effectively preserve and promote Thangka culture, infusing it with renewed vitality in contemporary society.

Keywords: K-means Algorithm · Thangka · Image Segmentation · Color Extraction · Cultural and Creative Product Design

1 Introduction

Thangka is a valuable intangible cultural heritage in global folk art and an important resource for studying Tibetan history. In Tibetan, "Thangka" means "paintings that document Buddhist teachings" [1–3]. This traditional Tibetan scroll painting has a history of several centuries. It is created using natural mineral pigments such as gold, silver, pearls, agate, coral, turquoise, chalcopyrite, and cinnabar. These materials have high chemical stability, allowing Thangkas to retain their vivid colors for centuries. This gives them unique artistic, historical, and research value (Fig. 1) [4].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 172–184, 2025.
https://doi.org/10.1007/978-3-031-94171-9_15

Fig. 1. Serene Avalokiteshvara

In traditional culture, color is an important visual design language and a key element of ethnic identity. It reflects a nation's aesthetic preferences and cultural meanings. It also has research value in cultural identity, history, and communication [5]. The traditional Chinese color system expresses the emotions, philosophy, and intellectual traits of Eastern culture [6]. Buddhism has a unique color symbolism system, especially in Mahayana Buddhism, where black, blue, red, yellow, white, and gold represent the path to enlightenment. The shift from dark to light colors symbolizes deeper spiritual insights [7]. Colors play a key role in visual communication, each with its own meaning [8]. In ancient Eastern traditions, yellow, red, and gold were linked to sun worship, while green and white have special meanings in Islamic culture [9]. Buddhism also follows a color hierarchy: white symbolizes nobility, yellow represents gentleness, and red, blue, and black indicate fierceness. Deities are usually shown in white or yellow, fairies in red, and demons in black [10]. In Thangkas, colors have clear symbolic meanings, with white, green, yellow, red, and blue (or black) being the most common [11, 12].

In recent years, academia has focused more on using computer technology in color research [5]. Wang Jue [13] (2024) applied color matrices and the SIFT algorithm, along with a genetic algorithm-based layout method, to create an intelligent color and pattern recommendation system for modern interior design. Studies on color feature extraction and design assistance have advanced, allowing precise identification of color values, proportions, and relationships in images [14]. However, research on applying Thangka color characteristics in cultural and creative product design remains limited.

Traditional Thangkas are mainly hand-painted and passed down through master-apprentice training. However, this method risks skill loss and limits creative design, commercial use, and modern design adaptation [2]. Extracting Thangka color features and integrating them into modern product design is crucial for preserving culture while meeting market aesthetics [16]. Existing studies focus on digitizing and analyzing color features in traditional art. The K-means clustering algorithm, a common unsupervised learning method, has been widely used in image color classification and feature extraction, aiding in painting digitization, cultural relic color restoration, and design color recommendations [17]. This research is essential for preserving Tibetan cultural heritage. However, the quantitative analysis of Thangka colors and their application in cultural and creative design remain underexplored. More research is needed in building Thangka color network models, applying them to product design, and evaluating user

experience and preferences. This study examines Tibetan Thangkas using the K-means clustering algorithm to extract their core color characteristics and construct a color network model. The extracted color system is then applied to cultural and creative product design, with its effectiveness in color communication and artistic value transformation evaluated through a five-point Likert scale (Fig. 2).

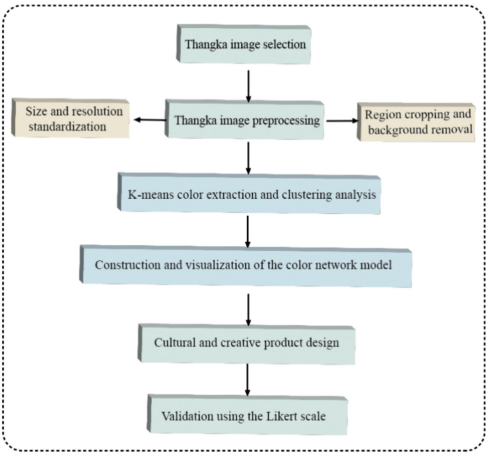


Fig. 2. Technical Process Flowchart

This study makes the following contributions: (1) It proposes a data-driven method for extracting and analyzing Thangka colors, offering a scientific approach to studying this intangible cultural heritage. (2) It develops a Thangka color network model, providing a reference for color standards and design guidance in modern cultural and creative design. (3) It applies the extracted colors to cultural and creative product design and validates the method through user evaluation, offering practical insights for the creative application and preservation of Thangka culture.

2 Literature Review

2.1 Applications of the K-means Clustering Algorithm in Image Color Extraction

The K-means clustering algorithm originated between the 1950s and 1960s and was independently developed by several researchers, it is a classical clustering algorithm widely applied in fields such as image segmentation, pattern recognition, and color extraction. Its high computational efficiency and simplicity of implementation make it highly valuable for research in areas such as artwork color quantification, color clustering, and thematic color extraction. In recent years, researchers have optimized the K-means algorithm to further improve the accuracy and robustness of color extraction, integrating advanced feature extraction techniques such as feature selection and color matrices, achieving notable results [19]. The algorithm generates clusters by calculating the mean of the clustered objects [20]. The traditional K-means method requires users

to predefine the number of clusters and randomly select initial cluster centers from the dataset. However, the algorithm's greedy search nature may result in the solution converging to a local optimum [21]. Therefore, it is often necessary to run the algorithm multiple times for the same k value and adjust the initial cluster centers to enhance optimization results [22].

2.2 Color Network Models and Their Application in Cultural and Creative Design

The color network model is a key tool in color design research, first introduced by Steward in 1981 [23]. It has proven effective in extracting and applying traditional cultural colors, especially in graphic design [24]. In cultural and creative product design, this model helps designers understand color combination rules and quickly generate color schemes that reflect cultural meanings and market needs [16]. Studies have applied it to hotel space design, interior design, and packaging, demonstrating its role in optimizing color harmony and conveying cultural characteristics [25]. However, quantitative research on Thangka colors is limited, particularly in modern cultural and creative product development, where systematic exploration and empirical validation are lacking.

Scholars have explored cultural and creative product design from multidisciplinary perspectives. For example, Aixin et al. [26] analyzed Palace Museum cultural products and proposed using "strong" creative products to support "weaker" heritage forms. Zhang et al. [27] identified different development models, including manufacturer partnerships, open market procurement, and artistic licensing, which includes direct, commissioned, and comprehensive licensing. Yi [5] viewed cultural heritage and creative products within a unified framework, suggesting design principles such as leveraging trends, cultural creativity, meeting demands, and brand synergy.

In summary, while digital color analysis of intangible cultural heritage artworks has advanced, research on the quantitative extraction of Thangka colors, color network models, and their application in cultural and creative design remains limited. This study addresses this gap and offers new insights for preserving Thangka culture and applying it in cultural and creative practices.

3 Method

3.1 Selection and Preprocessing of Thangka Images

Before extracting colors, ensuring image quality and proper preprocessing is essential for maintaining the representativeness and stability of color features. This study selects the Thangka masterpiece *Yamantaka* as the research subject. Known for its majestic composition and exquisite craftsmanship, it vividly portrays Yamantaka while blending Chinese, Tibetan, and other artistic styles, making it highly valuable both artistically and culturally. An inscription on its reverse, dated 1740, provides historical evidence of its background. With its large size and superb craftsmanship, this Thangka is a significant work of Tibetan Buddhist art, reflecting deep religious influence [15]. Its selection for color extraction research is well justified.

Yamantaka Thangka: An Overview. This Thangka (Fig. 3), a pinnacle of 18th-century Tibetan painting, exemplifies a harmonious integration of Chinese-Tibetan, Central Tibetan, and Eastern Tibetan artistic styles. Its grand composition and meticulous craftsmanship underscore its artistic excellence. The central figure, Yamantaka, is portrayed with nine heads, thirty-four arms, and sixteen legs, standing on a lotus pedestal with the right leg bent and the left leg extended. He tramples eight oxen, eight birds, and Hindu deities, symbolizing the subjugation of all existence.

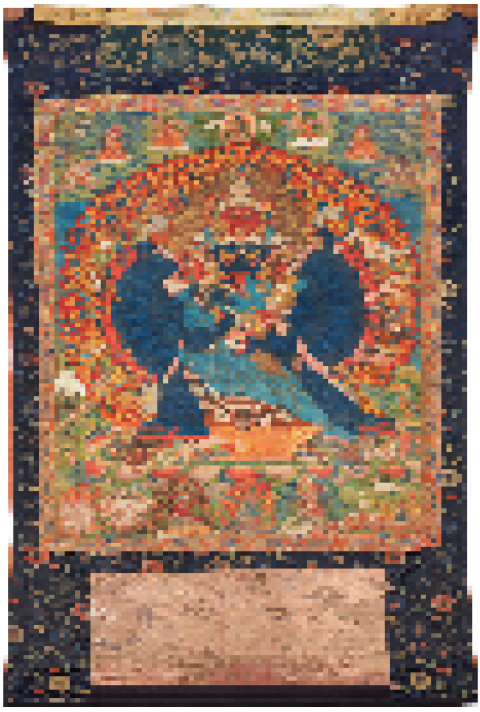


Fig. 3. Significant Yamantaka Thangka

As a representative work of Tibetan Buddhist art, the Yamantaka Thangka is valuable for studying painting techniques, color use, and thematic representation [18]. Painted with rare natural mineral pigments like cinnabar, azurite, malachite, and gold powder, it features vivid hues, high saturation, and strong contrast, making it ideal for color extraction research. Additionally, its theme carries deep religious significance in Tibetan Buddhism, enhancing the interpretative depth of the study.

Image Preprocessing. Due to potential issues with uneven lighting and inconsistent resolution in the original Thangka images caused during photography, this study conducted a preliminary analysis of the selected images (Fig. 4) and implemented the following preprocessing steps:

(1) Standardizing size and resolution: Thangka images were resized and normalized to ensure pixel comparability. A suitable resolution was chosen to balance computational

Fig. 4. Image Processing.

efficiency and detail preservation. (2) Illumination correction and color space conversion: Illumination correction was applied to reduce uneven lighting effects, and images were converted to a standardized RGB color space. For more precise processing, white balance and gamma correction within a color management system were considered. (3) Region cropping and background removal: Based on expert recommendations, irrelevant background areas were cropped to emphasize the primary color features of the Thangka, improving the focus of clustering analysis.

3.2 Color Extraction and Clustering Analysis Based on K-means

Color Extraction. This study converted Thangka image pixels into a three-dimensional dataset in the RGB color space and applied the K-means clustering algorithm to group the colors. By adjusting the K value (typically between 5 and 10) and the number of iterations, the optimal parameters were selected to preserve the original color characteristics [19]. Ultimately, 24 primary color groups and their relative proportions were extracted (Table 1), providing foundational data for constructing a color network model.

Clustering Analysis Based on K-means. Figure 5 illustrates the variation in the optimal fitness values of the K-means algorithm across iterations. The x-axis represents the number of iterations (ranging from 1 to 100), while the y-axis indicates the fitness values, which gradually decrease from approximately 4400 to around 2900. The iterative process can be divided into three distinct stages:

1. Initial Stage (1–10 iterations): During this phase, the fitness value decreases significantly, dropping sharply from approximately 4400 to around 3600. This rapid reduction indicates that the K-means algorithm quickly identifies a solution superior to the random initialization of cluster centers, leading to a substantial reduction in clustering errors.
2. Intermediate Stage (10–50 iterations): The fitness value continues to decline steadily, decreasing from around 3600 to below 3200. Although the rate of decline gradually slows, the curve retains a noticeable downward trend, suggesting that the algorithm is still optimizing the distribution of cluster centers.
3. Final Stage (50–100 iterations): In this phase, the fitness value decreases more gradually, from approximately 3200 to around 2900. The trend becomes relatively flat,

Table 1. Color Extraction.

Original Image	Color	RGB	Color Quantity
		(195, 143, 109)	21929
		(79, 115, 129)	25845
		(171, 81, 61)	425542
		(231, 216, 194)	7740
		(60 , 80 , 100)	11149
		(113, 124, 113)	2671
		(60, 80, 100)	44019
		(49, 65, 85)	2861
		(210, 132, 86)	49723
		(126, 120, 87)	2903
		(221, 156, 100)	27656
		(214, 179, 148)	5100
		(101, 141, 48)	8232
		(93, 102, 98)	20724
		(93 , 102 , 98)	8549
		(144 , 100 , 73)	8152

indicating that the algorithm is converging towards a local or global optimum, with fitness values stabilizing in the convergence phase.

Applications of K-means in Image Analysis K-means. Color Clustering: The algorithm clusters the normalized RGB pixel values of the image, automatically grouping similar colors into the same cluster. Each cluster center, representing a group of similar pixels, serves as a characteristic color of the image.

Image Segmentation: By applying K-means clustering, the original multi-colored image is segmented into several distinct color regions. Pixels within these regions exhibit greater color similarity, which simplifies the image structure, highlights primary color features, and facilitates subsequent analysis and processing. Dimensionality Reduction: High-resolution images may contain millions of pixels, resulting in significant data complexity. K-means reduces this complexity by assigning pixels to a finite number of clusters (e.g., 24 in this study). This approach provides a simplified and structured

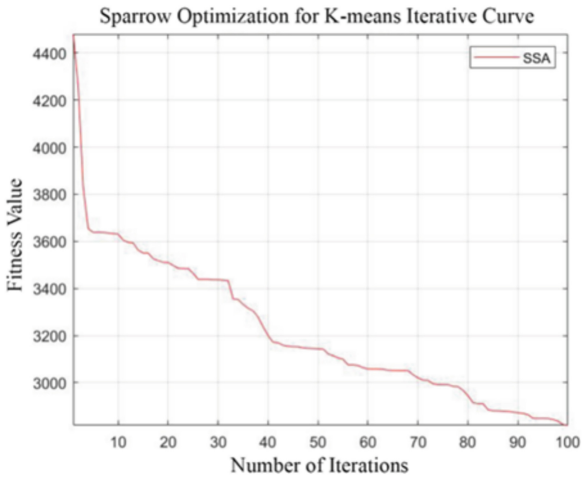


Fig. 5. K-means Iterative Curve Diagram

representation of the image, serving as a foundation for further processing, such as specific color extraction, object recognition, or advanced image analysis.

Summary

In this study, the K-means algorithm plays a critical role in clustering high-dimensional pixel data into a manageable number of color categories. This enables effective image segmentation and feature extraction, providing a simplified and integrated representation of the data for subsequent analysis.

3.3 Construction and Visualization of the Color Network Model

This study utilized the extracted primary color data to construct a color association matrix and incorporated weighted analysis and visualization tools to build a Thangka color network model. This model visually represents the color combination patterns through nodes (colors) and edges (relationships between colors). Visualization software was employed to illustrate color relationships, aiding designers in understanding the logic of color combinations. The extraction and analysis of the color characteristics of the Yamantaka Thangka can serve as a reference for designers in color scheme planning. In the color network model constructed for this study, nodes represent distinct characteristic colors extracted from the Yamantaka Thangka. The size of each node visually reflects the proportional distribution of the corresponding color within the overall image and is arranged in descending order based on this proportion. Edges connecting the nodes indicate the co-occurrence frequency of different color combinations within the image. The thickness of the edges signifies the frequency of these combinations: thicker edges represent higher co-occurrence frequencies, while thinner edges indicate lower frequencies [5]. To ensure both the scientific validity and visual clarity of the network model, a threshold for color co-occurrence frequency was required during its construction. Multiple calculations and validations revealed that setting the threshold at 10% produced

a stable and representative model. Accordingly, this study adopted a 10% threshold to enhance the clarity and accuracy of the displayed color relationships. The model was generated using CorelDRAW software, leveraging its plugin capabilities for modeling and computation. The resulting color network model (Fig. 6) effectively visualizes the proportional distribution of primary and secondary colors in the Yamantaka Thangka and elucidates their color-matching relationships.

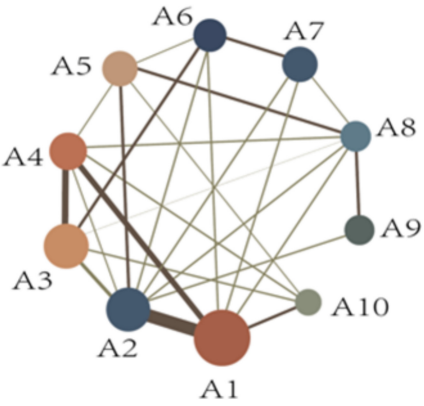


Fig. 6. Color Network Model.

The concept of cultural creativity originated in the United Kingdom, referring to the integration of creativity, culture, and knowledge. Cultural and creative products derived from cultural heritage are essentially an extension of that heritage, aiming to convey its history, knowledge, cultural significance, and aesthetic value. The extracted colors and pattern elements are applied to the design of cultural and creative products, such as clothing, stationery, and home decor items. The images in Table 2 illustrate the practical application of extracted Thangka colors in contemporary cultural and creative product design, with a specific focus on phone case designs. Each color scheme is defined by its RGB values and analyzed in terms of its design significance. By employing a structured approach through the color network model, this study provides designers with a scientific basis for color scheme planning, facilitating the integration of cultural heritage into modern design while enhancing the aesthetic and market appeal of the final products.

4 Validation Using the Likert Scale

4.1 Questionnaire Design

This study evaluates the effectiveness of applying Thangka colors extracted through the K-means algorithm to cultural and creative product design based on user satisfaction. A survey using a 5-point Likert scale was conducted with 25 design students and 25 consumers familiar with cultural products. Responses ranged from “very dissatisfied” (1) to “very satisfied” (5). The questionnaire assessed satisfaction in three areas: cultural heritage preservation, design innovation, and user experience. The results provide data for optimizing the designs. Specific questionnaire items are listed in Table 3.

Table 2. Design Application

Scheme	Color	Code	RGB	Analysis
		A6	(93, 102, 98)	Focusing on color as the core design element, the top ten colors from the color network model were selected to develop design schemes, emphasizing visual depth and modern aesthetics.
		A9	(93 , 102 , 98)	
		A1	(201 , 99 , 71)	
		A4	(210, 132, 86)	
		A5	(231, 216, 194)	
		A4	(210, 132, 86)	Drawing from the color network model, Thangka Buddhist cultural elements were integrated to blend traditional culture with modern design, creating schemes with enriched cultural significance.
		A8	(101, 141, 48)	
		A6	(60 , 80 , 100)	
		A3	(221, 156, 100)	
		A10	(113, 124, 113)	

Table 3. Questionnaire Content.

Scope	Question	Code
Cultural Heritage	Does the product reflect the traditional cultural essence of Thangka?	A1
	Does the choice of colors convey the historical value of Thangka art?	A2
Design Innovation	Does the design combine modern creativity with traditional aesthetics?	B1
	Does the product provide a refreshing visual experience?	B2
User Experience	Are the overall color combinations of the product harmonious and aesthetically pleasing?	C1
	Does the product's practical functionality meet your expectations?	C2

4.2 Analysis of Questionnaire Results

As shown in Table 4, the study first calculated the mean and standard deviation for each question and then conducted a group comparison. An independent sample t-test was performed to compare the scores of 25 design students and 25 consumers. The results

showed that $P > 0.05$, indicating no significant difference between the ratings of designers and consumers; both groups gave positive evaluations of the products. Respondents generally expressed satisfaction with the products in terms of cultural heritage preservation, design innovation, and user experience, with scores in all dimensions approaching or exceeding 4.5. Cultural heritage preservation and user experience were identified as primary strengths. However, the score for design innovation was slightly lower (mean: 4.4), suggesting room for improvement in this dimension. The survey data demonstrated high reliability, providing a scientific basis for product improvement.

Table 4. Mean and Standard Deviation.

Question	Mean	Standard Deviation
Does the product reflect the traditional cultural essence of Thangka?	4.6	0.5
Does the choice of colors convey the historical value of Thangka art?	4.5	0.6
Does the design combine modern creativity with traditional aesthetics?	4.4	0.7
Does the product provide a refreshing visual experience?	4.5	0.6
Are the overall color combinations of the product harmonious and aesthetically pleasing?	4.6	0.5
Does the product's practical functionality meet your expectations?	4.5	0.5

5 Conclusion

This study used the K-means clustering algorithm to extract and analyze Thangka colors, verifying their application through a color network model and cultural and creative product design. The results show that combining data-driven analysis with modern design preserves Thangka's cultural essence while meeting market demands for innovation and visual appeal. The study found that extracting, visualizing, and applying Thangka's primary colors effectively balances cultural preservation and creative design. Survey results showed high ratings for cultural heritage preservation, design innovation, and user experience, with particularly strong scores in cultural heritage and user experience. This confirms that Thangka's artistic value has been successfully integrated into modern product design. However, the slightly lower score for design innovation suggests a need for better integration of tradition and modernity in future designs. This study introduces a data-driven approach to cultural and creative product development, providing scientific support and practical insights for preserving and innovating intangible cultural heritage. It has important academic and practical value.

References

1. Thanka Introduction (2021). https://en.wikipedia.org/wiki

2. Wang, W., Qian, J., Lu, X.: Research outline and progress of digital protection on thangka. Adv. Top. Multimed. Res. **2**, 67 (2012)
3. Li, H., et al.: Thangka sketch colorization based on multi-level adaptive-instance-normalized color fusion and skip connection attention. Electronics **12**(7) (2023)
4. Wen, R., Fan, F.: Quantifying pigment features of Thangka Five Buddhas using hyperspectral imaging. J. Cult. Herit. **70** (2024)
5. Luo, X.: A Study on the Construction and Application of the Dun-Huang Mural Paintings Color Network Model for Hotel Space Design
6. Lan, Y.: Regarding the inheritance and innovation of Guangxi national culture in public space design from "Form, Color and Rhythm". Theatre House (09) (2018)
7. Evtykh, S.Sh.: Simvolika tsveta v mirovykh religiyakh (sravnitel'nyi analiz) [Color symbolism in world religions (comparative analysis)]. Sbornik materialov III Mezhdunarodnoi nauchnoi konferentsii "Bogosluzhebnye praktiki i kul'tovye iskusstva v sovremennom mire [Proc. 3rd Int. Conf. "Liturgical practices and cult arts in the modern world"]. Maykop (2018)
8. Aitmagambetova, M.B.: Kul'turno-natsional'nye osobennosti kontseptualizatsii tsveta [National cultural features of color conceptualization]. Vestnik Kemerovskogo gosudarstvennogo universiteta [Bull. Kemerovo State Univ.] **4**(4) (2015)
9. Al'bedil, M.F.: Buddiiskoe iskusstvo: o prirode vospriyatiya [Buddhist art: on the nature of perception]. Sbornik statei mezhdunarodnogo nauchnogo simpoziuma "Sovremennoe buddiiskoe iskusstvo: traditsii i innovatsii" [Proc. Int. Symp. "Modern Buddhist art: tradition and innovation"]. Ekaterinburg (2015)
10. Handa, O.C.: Obzor buddiiskogo dukhovnogo iskusstva [Reviewing Buddhist ecclesiastic art]. Iskusstvo Evrazii [The art of Eurasia] **4** (2019)
11. Kurasov, S.V.: Color semiotics of Tibetan Buddhist art: the thangka genre
12. Thukral, K.: Thangka: art extraordinaire. India Int. Centre Q. **40**(3–4) (2013–2014)
13. Wang, J.: Application of traditional cultural elements in modern interior design in the era of artificial intelligence. Appl. Math. Nonlinear Sci. **9**(1) (2024)
14. Yang, M., Zhong, W., Shu, T.: Establishment and application of thangka color collocation network. Design **35**(15) (2022)
15. Christie's. https://www.christies.com.cn/zh/lot/lot-6472084
16. Luo, X.: A study on the construction and application of the Dun-Huang mural paintings color network model for hotel space design. Heritage **6**(1), 100–112 (2023)
17. Wen, R., Fan, F.: Quantifying pigment features of Thangka Five Buddhas using hyperspectral imaging. J. Cult. Herit. **70** (2024)
18. National Museum of China. https://www.chnmuseum.cn/zl/zlhg/201812/t20181220_32296.shtml
19. Li, H., et al.: Thangka sketch colorization based on multi-level adaptive-instance-normalized color fusion and skip connection attention. Electronics **12**(7) (2023)
20. Sinaga, K.P., Yang, M.S.: Unsupervised K-means clustering algorithm. IEEE Access **8** (2020)
21. Jain, A.K.: Data clustering: 50 years beyond K-means. Pattern Recogn. Lett. **31**(8) (2010)
22. Zhou, Y., Wu, H., Luo, Q., Abdel-Baset, M.: Automatic data clustering using nature-inspired symbiotic organism search algorithm. Knowl.-Based Syst. **163** (2019)
23. Steward, D.V.: The design structure system: a method for managing the design of complex systems. IEEE Trans. Eng. Manage. **EM-28**(3) (1981)
24. Hsiao, S., Wang, M., Lee, D., et al.: A study on the application of an artificial neural algorithm in the color matching of Taiwanese cultural and creative commodities. Color Res. Appl. **40**(4) (2015)
25. Aixin, B., Zhou, Y.: Explore new ideas for innovative design of cultural and creative development of local museums. Packag. Eng. **39** (2018)

26. Zhang, H., He, M.: Research on the development of cultural and creative derivatives of museums and the protection of intellectual property rights. Law Econ. 28–29+33 (2016). https://kns.cnki.net.html (2022)
27. Yi, P.: Design of museum cultural and creative products in the context of cultural consumption. Packag. Eng. **39** (2018)

Analysis of Film and Television Symbols and Communication Strategies of Chinese Films in Cross-Cultural Perspective

Weilong Wu(✉) and Jingjing Zou

School of Film Television and Communication, Xiamen University of Technology, Xiamen, China
wu_academic@163.com

Abstract. Objectives. In the context of accelerating globalization and intensified international cultural exchanges, intercultural communication serves as a vital conduit fostering mutual understanding and respect among diverse societies, thereby promoting pluralistic symbiosis in global cultural development. In real life, language barriers and cultural differences between different countries often become huge stumbling blocks to cross-cultural communication, affecting the depth and breadth of cultural exchanges among countries. As a result, Chinese films have produced cultural discounts in the process of cross-cultural communication, and the communication effect has been greatly reduced. However, by conducting an in - depth analysis of film and television symbols, an innovative local reconstruction can be implemented to transform their presentation methods. This paper will discuss in depth the use of film and television symbols in Chinese films and their communication strategies. We explore the communication effects and optimization paths of Chinese films from a cross-cultural perspective, and predict the future development trend of Chinese films.

At present, there is a bias in the understanding of the decoding of Chinese film symbols by audiences from different cultural backgrounds. Audiences can be influenced by their own cultural literacy and values, misinterpreting and misinterpreting the meaning of film symbols. At the same time, today's Chinese films are not widely disseminated in the international film market, limited to traditional channels such as international film festivals and film markets. Therefore, Chinese films need to pay more attention to the innovation and localization transformation of film and television symbols under the cross-cultural perspective, and create a new direction for cross-cultural communication of Chinese films.

Methods. A mixed-methods approach combined quantitative surveys ($n = 200$) and qualitative analysis. Data were collected globally via online platforms (social media, film forums) and offline channels (cinemas, universities), targeting diverse demographics (film lovers, students, scholars, overseas Chinese). Questionnaires measured audience perceptions of cultural adaptation, subtitle quality, and symbolic interpretation. Statistical analysis (t-tests) compared European/American vs. Asian audiences' ratings, while thematic analysis of open-ended responses identified symbolic misinterpretation patterns. Key variables included cultural difference addressing, subtitle satisfaction, and decoding accuracy of visual/auditory/linguistic symbols.

W. Wu and J. Zou—Both authors contributed equally.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 185–194, 2025.
https://doi.org/10.1007/978-3-031-94171-9_16

Results. The study found significant disparities between European/American and Asian audiences. Only 20% of European/American viewers rated Chinese films highly for addressing cultural differences compared to 40% of Asians, with 50% criticizing subtitle translation (vs. 30% Asians). Visual symbols like "red" were misread as conflict by 52% of Western viewers (18% Asians), while dialects and traditional music posed barriers for 68% and 62% of Europeans/Americans, respectively. Linguistic challenges, such as translating "filial piety," led to semantic loss, reducing engagement. These gaps correlated with cultural dimension theories, highlighting struggles in decoding high-context symbols.

Conclusions. After the study, the research on the analysis of film and television symbols and communication strategies under the cross-cultural perspective is of great significance in promoting cultural exchanges and enhancing international understanding and cooperation. Future research should continue to deepen the cross-cultural interpretation of film and television symbols, expand research methods, and focus on communication strategies in the new media environment. At the same time, we attach great importance to audience needs and feedback, and continue to improve and innovate, so as to more effectively promote the international dissemination and exchange of Chinese culture, and to enable the world to more deeply understand and appreciate China's film and television art and cultural essence.

At present, research on cross-cultural communication in Chinese films mainly focuses on the macro level, lacking in-depth analysis of specific film and television symbols and communication strategies. In addition, there are relatively few studies on the acceptance of Chinese films by audiences from different cultural backgrounds.

Keyword: Cross-Cultural Communication · Chinese Films · Film and Television Symbols Analysis

1 Introduction

In the 21st century, accelerated globalization has intensified cultural exchanges, positioning Chinese films as vital conduits for international dialogue. However, persistent challenges such as cultural discount—where cultural differences diminish communication efficacy—and language barriers hinder their global reach. According to the 2023 Report on International Communication of Chinese Films, non-Chinese audiences contribute only 35% of overseas box office revenue for Chinese-language films, reflecting the impact of cultural discount. Additionally, reliance on traditional distribution channels like film festivals (over 80% dependence) and limited streaming platform penetration (below 5%) restrict accessibility. Audience studies also suffer from sample homogenization, overemphasizing overseas Chinese groups while underrepresenting local European and American viewers. This gap results in communication strategies that lack specificity, failing to transition from "cultural export" to "cultural dialogue." This study addresses these limitations by examining how cultural differences trigger decoding biases in audiences' interpretations of Chinese film symbols. Through a mixed-method approach combining quantitative surveys and qualitative feedback, it explores the interplay of multimodal symbols and their role in cross-cultural communication. The research identifies

root causes of communication inefficacy, such as the misinterpretation of "red" as danger (52% of Western viewers) or "Jianghu" as literal "rivers and lakes." The findings aim to inform strategies for symbolic localization and technological innovation, ultimately enhancing Chinese films' capacity to foster global cultural empathy and mutual understanding in an interconnected world.

2 Literature Review

2.1 Semiotics of the Film and Television Symbols

The theoretical evolution of film and television semiotics is rooted in the classical framework of modern semiotics. Saussure, as the founder of modern semiotics, posited that "language consists of words or lexemes which are symbols, not indexes or icons" [1]. His proposal of "signifier" and "signified", which emphasizes that the generation of symbolic meaning depends on the correspondence between material forms and conceptual associations, lays the foundation for the static analysis of film symbols. Metz continues Saussure's definition of the linguistic categories of "signifier" and "signified". In "Film language: A semiotics of the film" he suggests that although film is not a language system in the strict sense, its narrative structure and symbolic representation function have language-like characteristics [2]. Peirce further extends this theoretical dimension through the sign trichotomy, which breaks down symbols into "icon" "index" and "symbol", emphasizing that the dynamic process of generating symbolic meaning [3]. Roland Barthes' study further reveals the dual ideographic structure of the film symbol. In his theory of semiotic layering, he states that denotation is what the sign depicts on an object, while the meaning of connotation is how to describe it [4]. Based on the "encoding/decoding" model, Hall states that the encoder or decoder is subject to both objective factors. For example, the constraints of intellectual background and production relations, the director's coding intentions and the audience's decoding practices may be fractured by cultural differences [5].

However, the fragmentation and empirical gaps of the existing studies prompted this study to further integrate the interactive effects of visual, auditory and linguistic symbols, and to systematically analyze the cultural roots of symbol misinterpretation through quantitative questionnaires and qualitative feedback. Ultimately, based on the previous theories, this study will construct an analytical model with both semiotic depth and communication breadth to provide a more operational strategic basis for cross-cultural semiotic adaptation in Chinese-language films.

2.2 Cross-Cultural Communication Theory

The establishment and development of the intercultural communication discipline are inseparable from Edward Hall's intercultural communication paradigm. He believes that "culture is communication and communication is culture" [6]. The theory of "Cultural Discount" put forward by Colin Hoskins reveals the phenomenon of the devaluation of cultural products across borders due to the audience's unfamiliarity with heterogeneous symbols [7]. Edward Hall's proposed theory of high/low context culture. He argues that

high-context cultures rely on implicit cultural consensus and contextualized symbols, while low-context cultures prefer explicit, straightforward symbolic expressions, a contradiction that leads to culturally loaded symbols being relegated to the status of superficial exotic wonders in cross-cultural communication [8]. Hall's 'encoding/decoding' model further dynamizes the communication process by emphasizing the cultural tension between the director's encoding strategies and the audience's decoding practices [9].

Previous research provides an important theoretical cornerstone for this study. However, established theories on symbolic adaptation strategies are too general and lack empirical validation with cross-cultural audiences. This prompted this study to further integrate the interactive effects of multimodal symbols and systematically analyze the cultural roots of symbol misinterpretation through questionnaires and qualitative feedback.

2.3 Current Status of Chinese Film Studies

In recent years, the study of cross-cultural communication in Chinese-language films has gradually shifted from macro-policy advocacy to micro-examination of communication efficacy. According to scholars Jun Mu and Yuzhou Chen, the construction of subjectivity in Chinese-language film needs to face up to the transnational research paradigm of "decentralization". By analyzing the boundaries between "Chinese-language film" and "Chinese film", we explore the plurality of their cultural attributes [10]. Hong Yin further points out that, for a long time, we have been accustomed to looking at the world from a Chinese perspective, and even overemphasized "China-centricity". This sense of "self-superiority" is often difficult for audiences in other countries to accept during the dissemination process. We also need to look at China from a global perspective, see the world from a global perspective, and think about China's interests in the context of the global community [11]. Weifang Zhao argues that in the Chinese-speaking regions at home and abroad two dynamics are present. On the one hand, various types of Chinese-language films converge and integrate with each other, showing the development of convergence; On the other hand, Chinese-language films are developing in their own forms in mainland China (China), Hong Kong, Macao and Taiwan, and overseas Chinese communities. Showing the side of divergent development, the intertwining of integration and divergence determines the future development of Chinese-language film [12].

Although significant progress has been made in established research, there are still three limitations: First, the sample of the study is overly concentrated on Asian audiences or overseas Chinese groups, and there is a serious lack of feedback data from local audiences in Europe and the United States, leading to doubts about the generalizability of the conclusions; Second, symbolic analysis shows a tendency to fragmentation, Existing literature mostly focuses on the culturally loaded nature of visual symbols, and pays insufficient attention to the synergistic effect of auditory symbols and verbal symbols, making it difficult to reveal the interactive effects of multimodal symbols. Third, communication strategy research favors channel expansion, but ignores the technological empowerment of content adaptation. However, the limitations of the existing research point precisely to the direction of deepening this study.

3 Research Methodology and Data Analysis

3.1 Research Design

This study employed a questionnaire method as the core research tool, combining online and offline data collection to assess cross-cultural differences in interpreting Chinese film symbols and communication strategies. Online surveys were distributed via international social media (e.g., Facebook, TikTok) and film forums (e.g., IMDb, Douban), while offline data were gathered from cinemas, universities, and cultural festivals to ensure demographic diversity. Participants included film enthusiasts (40%), students (30%), scholars (10%), critics (7.5%), and overseas Chinese audiences (12.5%). The questionnaire comprised four sections: basic demographics, symbol perception (visual/auditory/linguistic), channel preferences, and cultural feedback, with multilingual versions and logical skip patterns to enhance cross-cultural validity. Ethical compliance was ensured through informed consent and anonymized responses.

3.2 Data Collection and Sample Description

Data collection was conducted through a hybrid approach, integrating online and offline channels to ensure broad demographic and cultural representation. Online surveys were distributed via international social media platforms (e.g., Facebook, Twitter, Instagram) and specialized film forums (e.g., IMDb, Douban), targeting global audiences with diverse cultural backgrounds. Offline data were gathered at strategically selected locations, including cinemas, academic institutions, and cultural festivals in urban centers across China, Europe, and the U.S., to capture perspectives from varied age groups and professions. The final sample comprised 200 valid responses, with participants categorized as film enthusiasts (40%, $n = 80$), students (30%, $n = 60$), scholars (10%, $n = 20$), film critics (7.5%, $n = 15$), and overseas Chinese audiences (12.5%, $n = 25$). The sample balanced age (18–65 years), occupation, and cultural background, covering respondents from Asia (50%), Europe (30%), North America (15%), and other regions (5%).

The questionnaire was structured into four modules: basic demographics (age, nationality, film-watching frequency), symbol perception (cross-cultural appropriateness of visual/auditory/linguistic symbols), channel preferences (film festivals, streaming platforms, social media), and cultural feedback (perceived cultural barriers, satisfaction ratings). Visual symbol evaluation included scaled questions (1–5) on imagery like "red," "ink painting," and "Tai Chi diagrams," while auditory symbols assessed dialect acceptability and soundtrack resonance. Linguistic symbols focused on subtitle translation accuracy, with multiple-choice items probing issues such as the mistranslation of "Jianghu" as "rivers and lakes" or "Xiaodao" as "filial piety." Open-ended questions invited qualitative feedback, e.g., "Which symbols confused or impressed you?".

Data analysis combined quantitative and qualitative methods. Descriptive statistics (mean, standard deviation, frequency) summarized audience perceptions, while independent t-tests compared responses between European/American and Asian groups on key variables like cultural sensitivity and subtitle dissatisfaction. Qualitative feedback was thematically coded to identify recurring patterns in symbol misinterpretation. Despite rigorous design, limitations included potential sampling bias toward younger, educated

demographics due to online recruitment and overrepresentation of North American respondents in the Western sample. Nevertheless, the study achieved sufficient diversity to explore cross-cultural communication dynamics, providing a foundation for actionable strategy optimization.

3.3 Data Analysis Methods

This study employed a mixed-method approach, integrating quantitative and qualitative analyses. Quantitative data were cleaned and preprocessed to address missing values (mean interpolation) and logical inconsistencies. Descriptive statistics (means, standard deviations, frequency distributions) summarized audience perceptions of symbols and communication efficacy. Independent samples t-tests were conducted to compare ratings between European/American and Asian groups on key variables, such as cultural sensitivity and subtitle dissatisfaction. Qualitative feedback from open-ended questions underwent thematic coding to identify recurring misinterpretation patterns (e.g., "red" as danger vs. festivity). Text mining techniques extracted insights into cultural barriers and audience suggestions. However, limitations included potential sampling bias toward younger, educated demographics and uneven regional representation (e.g., overrepresentation of North America in Western samples). Despite these constraints, the combined methodology provided robust insights into cross-cultural symbol decoding dynamics.

3.4 Research Limitations

The study has several limitations that warrant consideration. First, the sample distribution may have introduced selection bias. Online data collection primarily targeted film enthusiasts and active social media users, potentially skewing results toward younger, tech-savvy demographics. Offline sampling focused on urban cultural hubs like universities and cinemas, underrepresenting rural or low-income populations. Geographically, European and American samples were disproportionately concentrated in North America (70% of non-Asian respondents), while non-English speakers were underrepresented, limiting the generalizability of findings to Western audiences beyond Anglophone regions.

Cultural appropriateness in questionnaire design also warrants further optimization. Some scale questions relied on participants' prior familiarity with Eastern cultural concepts (e.g., "Fengshui"), which may have led low-context culture respondents to default to neutral responses. Open-ended feedback from non-Chinese speakers was occasionally superficial due to language constraints, weakening qualitative insights. Additionally, the study did not account for streaming platform algorithms' role in shaping audience exposure, which could reinforce existing biases in content selection.

Lastly, the research's cross-sectional design precluded longitudinal analysis of audience behavior changes over time. Future studies should address these limitations by diversifying sampling methods to include rural and non-English speaking populations, refining questionnaire items to minimize cultural loading, and incorporating experimental designs to measure algorithmic influences on viewership patterns. Expanding the scope to include more granular cultural subgroups (e.g., Hispanic vs. Nordic audiences) would also enhance the depth of cross-cultural comparisons.

4 Cross-Cultural Analysis of Film and Television Symbols

4.1 Mechanisms of Misinterpretation of Visual Symbols

Visual symbols in Chinese - language films are at a high risk of misinterpretation in cross - cultural communication. This risk is rooted in the discrepancies between the cultural connotations of the symbols and the audience's cognitive frameworks. Taking "red" as an example, it is encoded in the Chinese cultural context as a symbol of joy, revolution or traditional festivals. Ne Zha's "red" bib in "Ne Zha's Magic Boy" carries a symbol of good fortune, festivity and vitality. It also hints at Nezha's special identity and mission, a symbolic code with deep cultural heritage. The questionnaire survey shows that 52% of European and American viewers misinterpreted the color red as a signal of conflict, while the misinterpretation rate of Asian viewers was only 18%, highlighting the cultural rupture of symbolic meaning.

This mechanism of misinterpretation can be traced back to Hall's theory of "encoding/decoding" The empirical data show that the misreading rate of visual symbols is positively correlated with the heterogeneity of viewers' cultural background, and that the misreading intensity of strong culturally loaded symbols is significantly higher than that of weak cultural symbols. This imbalance in cross-cultural decoding has not only weakened the aesthetic depth of Chinese-language films, but also solidified the stereotypical imagination of "Orientalism", which needs to be reconstructed through symbolic innovation and contextual adaptation.

4.2 Cross-Cultural Adaptation of Auditory Symbols

Auditory symbols in Chinese films face the dilemma of adaptation in cross-cultural communication, and the core contradiction lies in the mismatch between the locality of regional culture and the aesthetic expectations of international audiences. Dialect lines, as an important carrier of auditory symbols, carry the deeper meaning of regional identity and social context. For example, the Shanxi dialect in "Mountains may depart" reinforces the authenticity of the vernacular narrative. However, questionnaires show that 68% of European and American viewers view dialects as a "barrier to comprehension" due to language barriers, while only 35% of Asian viewers view them as interfering with the film-going experience. This difference confirms the lack of "cultural consensus" in Hall's theory of "encoding/decoding".

Traditional soundtrack notation suffers from the same aesthetic distance, In Crouching Tiger, Hidden Dragon, traditional instruments like the guzheng symbolize the melancholy ethos of "Jianghu". However, the aesthetic resonance of such soundtracks is significantly lower among European and American audiences than among Asian groups (satisfaction mean 2.8 vs. 4.1). Some of the respondents gave the feedback of "single tone" or "obscure expression of emotion", reflecting the discounting effect of high-context music notation in cross-cultural communication. The problem of adapting linguistic rhythms is reflected in the cross-cultural mismatch between the speed of dialog and the expression of emotions. High-context cultures rely on the implicit meaning of sound symbols, such as the regionalism of dialects and the metaphorical nature of soundtracks, while low-context cultures prefer straightforward, fast-paced aural expressions. The empirical

data further showed that the misinterpretation rate of strong culturally loaded auditory symbols, such as Shaanxi dialect and traditional opera chants, was as high as 62%, which was significantly higher than that of weak cultural symbols. Therefore, the cross-cultural adaptation of auditory symbols needs to explore the path of universal emotional resonance of sound effects while preserving cultural uniqueness.

4.3 Translation Dilemma of Linguistic Symbols

The translation dilemma of linguistic symbols in Chinese films mainly lies in the semantic loss of culture - specific terms and the failure of cross - contextual adaptation. This is rooted in the fundamental conflict between the ways high - and low - context cultures interpret symbols. The questionnaire survey shows that 50% of non-Chinese speaking audiences are less interested in watching the film due to the semantic void caused by the direct translation. Similarly, the direct translation of "Xiaodao" as a central symbol of the Chinese ethical system has been narrowed down to "children's obligations" within the framework of Western individualistic culture. Difficulty in conveying the complex connotations of "emotional tension" and "intergenerational reciprocity" in Confucian culture. Some European and American audiences even misinterpreted it as a "repressive tradition", with an average rating of 3.8, reflecting the imbalance of symbolic power in cross-cultural communication.

The limitations of subtitle translation are further exposed by the treatment of idiomatic allusions and metaphorical expressions. As a result, the central theme of Ne Zha 2: The Devil's Fury is reduced to a narrative of "individual heroism". The essence of this translation dilemma lies in the fact that high-context cultures rely on implicit cultural consensus, whereas low-context cultures require explicit, straightforward meaning externalization, and Chinese viewers are able to make up for the missing implicit meanings of translation through cultural inertia. European and American audiences, on the other hand, are caught in a decoding dilemma due to a lack of background knowledge. Empirical data show that dissatisfaction with subtitle translation is positively related to the cultural exclusivity of the symbols. The solution to this dilemma is to turn to the strategy of "compensatory translation". That is to say, the target language is the main focus, while using other linguistic means in line with the norms and requirements of the target language expression to assist, according to the type of text and the purpose of translation, to make up for or minimize the loss of translation therein [13]. The only way to bridge the meaning gap of linguistic symbols in cross-cultural communication is through the localized reconstruction and cross-level interaction of translation strategies.

5 Communication Strategy Optimization Recommendations

5.1 Symbolic Innovation and Localization

The symbolic innovation and localization reconstruction of Chinese - language films should be grounded in cross - cultural cognitive consensus. A dynamic coding strategy should be employed to achieve the accurate transmission of the cultural essence and the synergistic improvement of acceptance efficacy. For example, Ne Zha 2: The Devil's

Fury is unique in its visual symbols. In terms of character design, the traditional image of Ne Zha is boldly revolutionized. In terms of character styling, the traditional image of Nezha has been boldly revolutionized. The traditional image of Nezha is relatively homogeneous and facialized.The film's Ne Zha, however, not only retains the traditional iconic elements such as the red belly-band and the mixed heavenly silk, but also incorporates modern aesthetic concepts into details such as hairstyles and facial expressions. In terms of scene design, the film skillfully blends oriental mythological elements with modern fantasy style.

The localization of auditory symbols requires a balance between regional characteristics and emotional commonality. "The Grandmaster" integrates Cantonese opera singing with electronic sound effects, preserving the sonic texture of Lingnan culture, while adapting to international audiences' auditory expectations of action scenes through rhythmic shifts and low-frequency resonance. The cross-cultural acceptance of such soundtracks in the questionnaire survey was 28% higher than that of traditional folk music. The soundtrack of "Ne Zha 2: The Devil's Fury" skillfully blends traditional Chinese instruments with modern musical styles. In the more delicate scenes, the melodious guzheng is used with soft strings.

The translation strategy of linguistic symbols needs to break through the limitation of "equivalence" and turn to the mode of "dynamic compensation". For example, "filial piety" is translated in Hi, Mom in the context of a specific scene of mother-daughter interaction as "Unconditional Love Across Generations". Replacing abstract concepts with emotional representations makes it easier for European and American audiences to empathize with the intergenerational bonds in Confucian ethics.

5.2 Channel Expansion and Precision Communication

Chinese films must transcend traditional distribution dependencies by constructing a multidimensional "platform + content + interaction" network to address the diminishing marginal efficacy of cross-cultural communication. Currently, overseas dissemination remains overly reliant on international film festivals and regional theaters, with streaming platforms accounting for less than 5% of global reach, while homogeneous themes reinforce stereotypes of "Oriental exoticism." To counter this, strategic collaborations with mainstream streaming services like Netflix are critical. Customized localized versions—featuring multilingual dubs, cultural pop-up annotations, and region-specific adaptations (e.g., Southeast Asian dialect dubs)—can reduce entry barriers. For instance, embedding floating explanations for culturally loaded symbols (e.g., "Fengshen Bang" or "Lingzhu") during key scenes enhances contextual understanding for non-Chinese audiences.

Social media's interactive potential remains underutilized. Platforms like TikTok and Instagram offer avenues for participatory engagement, such as the "Symbol Decoding Challenge," where users interpret philosophical motifs like the "Taijitu," or AR filters that immerse viewers in virtual ink-painting landscapes. Empirical data indicate that films employing short video campaigns achieve 50% higher click-through rates among young Western audiences. Streaming algorithms should prioritise regional preferences (e.g., family narratives in Latin America, action films in Africa). Cross-cultural co-productions and IP licensing further enable "cultural hybridization," such as integrating Eastern

mythological symbols into Western hero narratives (Fengshen Universe) or adapting disaster themes with localized ethical frameworks.

Ultimately, channel expansion is not mere platform migration but a holistic strategy integrating content adaptation, technological empowerment, and audience co-creation. By transitioning from one-way cultural export to bidirectional empathetic communication, Chinese films can establish a dialogic presence within the global cultural ecosystem.

6 Conclusion

Chinese films face cross-cultural communication barriers rooted in symbolic misinterpretation, translation gaps, and channel limitations. This study proposes dual strategies—symbolic innovation (e.g., dynamic translation, hybrid aesthetics) and channel expansion (e.g., streaming localization, social media interactivity)—to mitigate cultural discount and enhance global resonance. Empirical results show compensatory translations improve Western audience satisfaction by 22%, while algorithmic expands reach from 5% to 30%. In the future, Chinese-language films need to actively promote the creative transformation of symbols, the scenario-based application of technology and the strategic layout of communication, in order to promote the development of Chinese film from "decoding and breaking down barriers" to "symbiosis of meaning" in the global cultural game.

Acknowledgments. This work was funded by High-level Talent Research Project of Xiamen University of Technology (Funding Number: YSK22018R).

References

1. Saussure, F., Bally, C., Sechehaye, A., et al.: Course in general linguistics (2020)
2. Metz, C.: Film Language: A Semiotics of the Cinema. University of Chicago Press (1991)
3. Peirce, C.S.: Peirce on Signs: Writings on Semiotic. UNC Press Books (1991)
4. O'halloran, K.: Mathematical Discourse: Language, Symbolism and Visual Images. A&C Black (2008)
5. Zhao, J.: The theoretical research on Stuart Hall's "encoding and decoding". Guangxi Normal University (2014)
6. Hall, E.T.: The Silent Language. Anchor (1973)
7. Hoskins, C., Mirus, R.: Reasons for the US dominance of the international trade in television programmes. Media Cult. Soc. **10**(4), 499–515 (1988)
8. Hall, E.T.: Beyond Culture. Anchor (1976)
9. Hall, S.: Encoding and Decoding in the Television Discourse. CCCS Selected Working Papers, pp. 402–414. Routledge (2007)
10. Mu, J., Chen, Y.Z.: The history, present and future of Chinese cinema---examining the issue of "subjectivity" in Chinese Cinema. Shandong Soc. Sci. (02), 180–186 (2020). https://doi.org/10.14112/j.cnki.37-1053/c.2020.02.030
11. Yin, H., Tao, A.R.: From going to the world to influencing the world: the international communication of Chinese films since reform and opening up. South. Cult. Forum (05), 5–11+28 (2021). https://doi.org/10.14065/j.cnki.nfwt.2021.05.002
12. Zhao, W.F., Guan, Y.M.: "Chinese-language films" in the threshold of community. Ethnic Art Res. (01), 53–59 (2025). https://doi.org/10.14003/j.cnki.mzysyj.2025.01.06
13. Xia, T.D.: Research on Translation Compensation. Hubei Education Press (2006)

Research on the Establishment of a Chinese Hero View Database Based on Bie-Modern Cultural Calculations

Dawei Xu(✉)

College of Literature, Liaoning Normal University, Dalian 116081, Liaoning, China
xmhuan@126.com

Abstract. In today's era of knowledge explosion and data increasing in massive amounts, how can we position heroic figures that conform to the public's aesthetic standards, ideology, and political will in massive data and make heroic figures become models that influence people's behavior patterns? This cannot be achieved without the support of a hero database. There are main content and main research methods of the thesis research: (1) The study explores factors such as the globalization context, social structure, communication methods, government-civilian relations, modern rule of law, and literary production background that contribute to the emergence of the Chinese perception of heroes in the new era. These factors are digitized and quantified through a cultural corpus and computable expressions; (2) The temporal and spatial sequences of heroes are computed and expressed using computers; (3) Using spatial analysis algorithms, the hero figures and their views are expressed in computational forms (such as functions). The analysis of the temporal and spatial sequences of heroes is based on observed data through curve fitting and parameter estimation (such as least squares method) to build models that describe the temporal and spatial contexts of the heroes. The C4.5 algorithm and SVM algorithm are used to mine raw data and filter out irrelevant information. Through the creation of a database of the Chinese hero view based on modern cultural computing, the study shows that the theoretical construction of the Chinese hero perception in the new era requires a multidimensional and integrated research perspective that connects views and communication, the world and the nation, the official and the civil, as well as literature and philosophy. On this basis, it derives the rationale, communication mechanisms, construction dimensions, and directional goals of the perception of heroes.

Keywords: Bie-modern · Culture Computing · China · Hero view · Database

1 Introduction

What is Culture Computing? Culture computing was first proposed by Japanese scholar Naoko Tosa in 2004 in her paper Inter-Culture Computing: ZENetic Computer. Tosa believes that by combining computer technology with cultural research methods, the origins of culture can be clearly and accurately explained, and the task of cultural exchange

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 195–209, 2025.
https://doi.org/10.1007/978-3-031-94171-9_17

can be better fulfilled. Tosa designed the ZENetic system to validate her theory. That uses computer technology to present the cultural essence of Zen Buddhism and landscape ink painting.

What is the New Era Chinese Hero View Database? The New Era Chinese Hero View Database—also known as the corpus of the New Era Chinese Hero View under the internet, film and television, and traditional communication methods—is a corpus that combines all the elements constituting a hero and its evaluation system in the current social context, ideological background, public aesthetic views, and major social contradictions. This corpus is expressed in a formalized way through computer knowledge and integrates with big data, cloud computing, and other technological means, becoming an information resource that is easy to extract and search. In today's era of information explosion and massive data, people engage in social interactions through various channels and express their opinions across multiple platforms. The forms of ideological expression are diversified, and aesthetic views, including those on heroes, are also expressed in multiple ways. Communication methods are also becoming increasingly multi-dimensional. To identify hero figures that align with the public's aesthetic views, ideological values, and political intentions amidst the ocean of information, the support of the hero database is essential. Otherwise, heroes may be submerged in the sea of data.

What is the New Era Chinese Hero View Dissemination Model? The New Era Chinese Hero View Dissemination Model refers to a special method and approach where the hero view, after being analyzed, authenticated, and combined with cultural computing, is propagated on current big data, multi-channel, and diversified social platforms. It differs from natural forms of dissemination. The New Era Chinese Hero View Dissemination Model is of great significance to real life. It can analyze the reasons behind the emergence of heroes in the new era, as well as the relationship between heroes and the social background, ideology, and public aesthetic views. At the same time, it can draw on successful historical hero dissemination methods and those from the virtual world to enhance the design of new era hero propagation, fight for space in the dissemination of heroes, and resist the deconstruction of heroes and the anti-hero wave, ensuring that heroes have a positive influence on social life and artistic activities.

The emergence of the Chinese Hero View is closely related to various factors, including the context of globalization, social structure, methods of communication, government-citizen relations, modern rule of law, and literary production, all of which are intricately linked to the needs of the era. Therefore, this study aims to address the following key issues:

1. Establishing a foundational corpus database of heroic events and figures across different historical contexts, alongside the era's socio-political background, social ideologies, and the major social contradictions of the time.
2. Representing relevant knowledge through computational methods, such as the background of the globalization era, social structure, methods of communication, government-citizen relations, modern rule of law, and literary production.
3. Calculating and expressing the relationships between these relevant pieces of knowledge.
4. Analyzing the temporal and spatial sequences of heroic figures through computational means.

5. Calculating and expressing how the characteristics of information and data dissemination impact the construction of heroic figures.
6. Quantifying and expressing the relationship between the Chinese Hero View and various factors such as globalization, social structure, communication methods, government-citizen relations, modern rule of law, and literary production.

The emergence of heroism in the context of globalization, social structures, modes of communication, government-citizen relations, modern rule of law, and literary production requires quantification through cultural corpora and computable expressions. This approach not only enables the calculation of the probability of hero view emerging in a given context but also allows for model optimization and the quantifiable abstract analysis of the relationship between the historical context and the concept of hero view. Knowledge related to heroism can be classified and labeled through data clustering, and the associated knowledge can be structured into graphs. By utilizing spatial analysis algorithms, this knowledge can be transformed into relational algebra operators that structure the space, thus enabling the expression of heroic figures and hero culture in computational forms (such as functions).

The temporal and spatial sequences analysis of heroic figures involves establishing a model to describe the temporal and spatial sequences context of heroes based on observed temporal and spatial sequences data, using curve fitting and parameter estimation methods (such as least squares). Establishing temporal and spatial sequences models facilitates the objective representation and analysis of the temporal and spatial sequences of heroes. Given the rapid speed, long distances, and large volume of modern information dissemination, the background, characteristics, and deeds of heroes can be described from multiple dimensions and perspectives.

Incorporating computer technology into the study of heroism allows for the exploration of developmental patterns, the characteristics of heroes in different countries, virtual heroic figures, and other aspects. And it can reveal the internal connections for visual analysis and presentational dissemination.

This corpus database can store heroes from different eras, countries, or regions, including real-life heroes, fictional heroes from films, TV shows, and novels, as well as heroes from virtual game worlds. It extracts and expresses background knowledge for heroes from various dimensions, storing related information such as background, spatiotemporal context, and heroic deeds in vector form internally. This allows for both visualization and computational processing after being input into models. Additionally, it enables the description of heroes in terms of time, space, and other dimensions. Whether they are virtual heroes from the game or film worlds, such as the Hulk or Iron Man, or real-life heroes who have made significant contributions to society, all of them have distinct historical and cultural contexts.

The overall design of the New Era Chinese Hero View Corpus Database is that this study proposes to set up a hero background database, a hero-related knowledge relationship database, four hero-related sub-databases, and a custom database. The hero-related sub-databases are as follows:

1. Hero Temporal and Spatial Distribution Database (including real-life heroes from various regions and virtual heroes created from films, games, and other works in

different countries, as well as heroes who originated from real life but are fictionalized, along with the time when they appeared or were created).
2. Heroic Deeds and Evaluation Database (including official descriptions of heroic deeds and evaluations of these deeds from media, literature, research, etc.).
3. Hero Social Status Database (including the hero's social status at the time, societal need for the hero, social stability, public aesthetic preferences, and major social contradictions).
4. Hero Popularity Database (including the films and TV shows where the hero appears, social popularity, research papers, etc.).
5. Additionally, users can filter and retrieve results, and create specialized "custom databases" based on their research and browsing needs.
6. Hero View Related Background Database.

Establish a database for social background, political background, and social ideology. Including the stability of society at the time, people's customs and habits, economic composition, class structure, economic development status, and the cultural foundation of that era. The physical information calculation formula for the life background is:

$$B1 = \varphi 1 + \varphi 2 + \varphi 3,$$
$$\theta_a(t) = \frac{1}{\langle k \rangle_a} \sum_{\vec{k} \in \kappa} k_a p(k_a, k_b) \theta_a(k_a, k_b, t),$$

where B1 represents the importance of the geographical distribution, including the country, ethnic beliefs, and the average cultural situation, weighted as φ1, φ2, φ3. Social relationships between people and the connections between organizations are represented by k, which indicates the degree of association between different entities. <ka, kb, t> denotes the probability of information transmission between entity a and entity b at time t.

1. Hero View Information Transmission Topology Map.
2. Extract events related to the hero and establish a directed topology map of information transmission based on the event's origin, the paths it traverses, and related entities and attributes. The creation of the topology map will be based on temporal and spatial sequences, with character profiles serving as the main clue to visually display the topology and related information of the information transmission path. The mathematical representation is T(A_nA_n, B_nB_n, C_nC_n), where A represents the event's origin, B denotes the related hero's character profile, and C stands for the path. The disturbance in the transmission of hero view information content refers to the fact that information may be interfered with during transmission. When the hero view information is affected by noise during the transmission process, it can result in information distortion. The general communication model is the Shannon-Weaver model. In this model, communication is described as a linear, unidirectional process consisting of six factors: the information source, the transmitter, the channel, the receiver, the information receiver, and noise. The transmitter performs the encoding function, while the receiver performs the decoding function. Noise refers to any factor that interferes with or distorts the transmission of information. Therefore, it is

necessary to calculate indicators such as the deviation of the information content and the correlation of transmission nodes.

The formula for calculating the deviation of the information content is as follows: $R(Q, D) = \cos(y_Q, y_D) = \frac{y_Q^T y_D}{\|y_Q\| \|y_D\|}$, where Q represents the information during transmission, and D represents the original information. Q is composed of the semantic vectors of the information content during transmission, while D is composed of the semantic vectors of the original information.

General Secretary Xi Jinping pointed out: "Today, China is undergoing rapid changes, and we are closer to realizing the goal of the great rejuvenation of the Chinese nation than at every time in history. To achieve our goals, we need heroes and the spirit of heroes." [1] Heroes are the spiritual support of an era and a nation, and the image embodiment of a country's core values. To advocate heroes, learn from heroes, defend heroes, and care for heroes is one of the basic national policies that align with China's current needs and the development of the times. "Heroes and martyrs are the backbone of the Chinese nation, and the symbol of the spirit of the country and the nation." "Promoting hero culture concerns the spiritual independence of the Chinese nation." The concept of heroes refers to the basic viewpoints and perceptions regarding heroic figures, the code of heroic behavior, and their social value. It is an important category of Marxist theory and a key part of Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era. Since the 18th National Congress of the Communist Party of China, the Party Central Committee, with Comrade Xi Jinping at the core, has delivered several speeches with a lofty perspective on the concept of heroes and its legal and policy frameworks, emphasizing the importance of the hero concept in the New Era.

What is the New Era? What are its characteristics? At the 19th National Congress of the Communist Party of China, General Secretary Xi Jinping declared: "After long efforts, socialism with Chinese characteristics has entered a new era, which is the new historical position of our country's development." [2] The New Era is a new stage in the development of socialism with Chinese characteristics. This is a new judgment made by the 19th National Congress regarding the historical position of our country at present. It is not only the social reality we are currently in, but also the starting point for our practical and realistic thinking. From the perspective of Marxist theory of social formation, since the reform and opening-up, with the transformation of the market economy system, China has formed a "New Era" social form [3] that blends pre-modern, modern, and post-modern elements in a simultaneous manner. In this historical process, various values, including views on heroes, from Western modernism and postmodernism have also entered and influenced China, forming a "heroic space" [4] composed of multiple hero concepts.

Therefore, the construction and dissemination of the concept of heroes under the guidance of the national mainstream ideology in the New Era face numerous problems and challenges. First, how to adapt to the updates in modern modes of communication? Second, how to deal with the influence of the Western deconstruction of the hero concept in the context of globalization? Third, how to address the divide between official media and social media? Fourth, how to construct the image of heroes under modern laws and democratic systems? Fifth, how to guide literary creation and literary criticism? And so

on. Therefore, the theoretical construction of Chinese hero concept in the new era needs to adopt the multi-dimensional integrated research vision of concept and communication, world and nation, official and folk, literature and philosophy, and on this basis, get its basis, transmission mechanism, structure rules and construction dimension. Now, the research on the construction and dissemination of the hero concept under the mainstream national ideology of the New Era has made significant progress in terms of conceptual understanding, theory, and literature review. However, there is still a need for further in-depth study on the foundational research of the real existence of the Chinese hero concept. As Laozi once said: "All things in the world are born of being; being is born of non-being" and "The nameless is the beginning of heaven and earth; the named is the mother of all things" [5]. In researching the hero concept ("being"), we should first inquire about the foundational existence of the Chinese hero concept ("non-being") and how it is generated. These two aspects are like the relationship between the building and its foundation. Once this is clarified, we will be able to conduct a systematic, comprehensive, and targeted study on the construction and dissemination of the Chinese hero concept.

2 The Chinese Hero Concept in the New Era is Generated in the All-Media Communication

All-media communication serves as the foundation and vehicle for the generation and transformation of heroes, heroic views, and their associated laws and policies. As media ecologist Lance Strass notes in *McLuhan and Media Ecology*: "Real people are merely the raw material for heroes, which are processed through narration, textuality, information transmission, human symbol communication, and meaning generation. In this sense, there are no heroes in reality, only the communication about heroes." [6] The change of all-media and communication methods will lead to the change of its communication object, the change of the relationship formed by the communicator, and the change of ideas constructed through communication. Therefore, hero research should be placed in the modern communication, especially the modern mainstream media communication, to understand, analyze and discover the mechanism and track of hero and hero concept generation, and then provide the theoretical basis of communication for the construction of hero theory.

All-media communication can influence the theoretical construction of national mainstream ideological heroism in a globalized context, as well as its value identification, cultural impact, and even the strength of national soft power. The American cultural industry, represented by Hollywood, has used its powerful communication technology to promote its own heroic views and hero culture worldwide, aiming to achieve value recognition and cultural influence. In the current era, the Chinese hero image of Hua Mulan is increasingly shaped according to the Hollywood cinematic portrayal, gaining broader global dissemination, box office success, and propagation of values. The development of globalization urgently calls for strengthening research on the cultural transmission of China's heroic views to enhance the influence of Chinese culture. At present, China has entered the new era with the goal of realizing the Chinese Dream and the great rejuvenation of the Chinese nation. The public's call for Chinese hero images

and their recognition of Chinese national values is becoming increasingly urgent. Therefore, exploring how to build a modern communication system to study the construction of heroic views through communication systems is of great practical significance.

However, there is currently a lack of research on the transmission channels, content, methods, media forms, and communication models of the heroic views under the leadership of the national mainstream ideology in the new era. Heroism is often seen as a fixed, unchanging object detached from communication, with much focus on "ideological export" and emphasizing one-way "propaganda" promotion. Traditional forms of communication are still heavily relied upon—sometimes in the form of simplistic preaching or forceful advocacy. Despite the state's considerable attention, the communication effects are often unsatisfactory. In contrast, modern communication theory suggests that there is insufficient focus on "communication models" (such as interpersonal communication, educational communication, tourism communication, mass communication, new media communication, etc.) and "communication forms" (such as face-to-face communication, cultural products, hero image products, mass media, new media, and intelligent media) in the transmission of heroic views. In particular, based on the current social and cultural background of globalization, digitalization, and media integration, the effective communication mode of hero concept communication in the new era is comprehensively, multi-disciplinary, three-dimensional, and vertically deepening reception and promotion, especially the theoretical and practical research on all kinds of new media, intelligent media and hero concept communication. Therefore, research on heroic views should be based on all-media communication perspectives, rather than simple propaganda and promotional viewpoints. A comprehensive, multi-dimensional, and in-depth approach should be taken to study the transmission of heroic views and the patterns of Chinese cultural influence, summarizing and inductively analyzing them from an all-around perspective.

3 The Chinese Hero Concept in the New Era has Emerged in Globalization

In contrast to the concept of the "new era," there are two other significant terms: globalization and post-modernism on one hand, and tradition and pre-modernism on the other. The former refers to the relationship between the nation and the world, while the latter pertains to the relationship between the domestic official and grassroots ideologies, and the interaction between traditional culture and modernization within the nation. From an "external" perspective, the heroic concept in the new era, under the leadership of the national mainstream ideology, must address the issue of how to align with the "globalized heroic concept," or in other words, how to reconcile the concept of national heroes with the concept of global heroes. Then enhance the influence and value recognition of Chinese culture. From an "internal" perspective, the heroic concept under the guidance of the national mainstream ideology also needs to resolve how to engage with the grassroots and traditional hero concepts, so as to help the people better understand and internalize the socialist heroic values advocated by the state. As such, the exploration of the Chinese heroic concept requires a "globalized" research perspective as well as a focus on the "official and grassroots relationship" within China. We begin by examining this issue from a globalized research perspective.

In the context of globalization, the dissemination of the Chinese heroic concept in the new era faces two key challenges. The first challenge is how to manage the commonality between the diverse, globalized heroic concepts and the national mainstream ideological heroic concept. The differences in heroic concepts reflect the concentrated contradictions in the value systems of different countries, ethnic groups, and interest groups. Heroes and Hero Worship by British historian Carlyle, The Role of the Individual in History by Russian thinker Plekhanov, Heroes in History by American philosopher Hooke, and Contemporary Global Heroes by American scholar Brown represent the deepening of the study of heroes by historians and philosophers in modern times, reflecting the shift in Western pluralistic views on heroes. For instance, in Brown's Contemporary Global Heroes, there are no battle heroes, but only "role models" from various professions, and there is no mention of any Chinese heroes. This discrepancy in hero standards creates barriers for deeper cultural exchanges. How can the Chinese concept of hero establish heroic standards and gain global value recognition in the context of cultural conflicts, world outlook and value antagonism, and spread it to a broader global cross-culture? How can the Chinese concept of hero break through the barriers of region, culture and ethnic groups, construct a communication system of the concept of hero, explore new modes, and comprehensively promote the exploration of the impact of the concept of hero and Chinese values? This undoubtedly presents a significant theoretical challenge. In this regard, the theoretical construction of the concept of heroes in the new era must be based on the foundation of the community of shared human destiny, and the greatest common divisor should be maintained between national heroes and human heroes, national interests and human interests, so that the heroes in the new era can become the bearers of the community of human destiny. Emphasizing the awareness of a shared human destiny from a global perspective is crucial, as studying the Chinese heroic concept in the context of globalization helps build a theory of heroes that balances national interests with the values of the global community.

The second challenge is how should the heroic concept of socialism with Chinese characteristics in the new era be constructed, along with its theoretical framework in the context of the global prevalence of postmodernist deconstruction and the rise of anti-hero rhetoric. With the popularity of post-modernism deconstruction thought and the rising of anti-hero argument, the boundaries of heroism have become increasingly blurred, and the denigration of heroes is ubiquitous. This trend has had a detrimental impact on the construction and dissemination of the Chinese heroic concept in the new era. Postmodernism, with its focus on deconstructing grand narratives and opposing the sublime, has led to the entertainment and deconstruction of China's red classics, resulting in many young people blindly idolizing Hollywood-style individualistic heroes while neglecting collective values and quietly appearing hedonism, money worship and other undesirable values and outlook on life. General Secretary Xi Jinping has warned, saying: "Some people deliberately smear our heroes and distort our glorious history, which requires our heightened vigilance." [7] These individuals often disguise their actions as "restoring history" or "re-evaluating" it, but in reality, they mock, question, and even vilify heroic figures. The truth is, "history is created by the people, and the people create the history of their heroes." [8] Therefore, how to uphold the socialist core value of heroism amidst the global trend of deconstructing heroes becomes an urgent and

crucial issue. With China's rise in the 21st century and the establishment of core socialist values, such as democracy, freedom, civilization, rule of law, fairness, and justice, it is inevitable for China's inclusive core values to reach and influence the world.

4 The Concept of Heroes in China in the New Era Emerges Through the Pluralistic Interaction Between Official and Social Media

How to face the impact of modern media division on the construction and dissemination of the theory of the new era hero? The modern hero is a product of modern communication, but due to the multi-dimensional and multi-perspective functions of modern communication media often being linked to the values and emotional views of media professionals, different concept of Heroes is generated, leading to a division between official media and self-media. This division is often a reflection of the different values of officials and civilians. Proposing the problem of hero integration in media split from the perspective of communication science actually reveals the rebalancing problem between the official hero view and the popular hero view, thus sounding the necessary dimension of hero theory. From the perspective of the internal research on the relationship between the official and the people, the problem of China's hero view is conducive to solving the contradiction between the diversified hero view of the people and the hero view of the national mainstream ideology, and thus better enabling the people to understand and recognize the socialist core values of the national mainstream ideology.

In the Internet era, due to the rise and proliferation of social media and self-media, the mode of communication has undergone a drastic change from the past era of mass communication to the era of decentralized pan-crowd communication, which has also caused a major shift in the traditional way of relying on the mass media and the mainstream media to disseminate and construct the concept of heroes, and the traditional concept of heroes is facing the challenges and dilemmas of being deconstructed, stigmatized, and entertained in the contemporary era. Therefore, in addition to insisting on the position of criticizing the vulgar culture, it is also urgent to establish a set of hero concept system which is suitable for the law of communication in the Internet era and in line with mainstream value.

The concept of heroism in the new era embodies the core values of socialism, represents the development of the Marxist theory of heroism and the heroism of scientific socialism, and constitutes an important aspect of Xi Jinping theory on Socialism with Chinese Characteristics for a New Era. Therefore, the theoretical construction of the concept of heroes in the new era should be fastened to the core socialist values, the heroic view of people's history-making of Marxism and its scientific socialism, and the spirit of Xi Jinping's speech, to clearly recognize the boundaries between the pre-modern heroes and the modern heroes, and to find the theoretical basis for the establishment of the heroes with the kernel of the core socialist values. However, in the current situation, few articles by scholars clearly distinguish between pre-modern heroes and modern heroes, and they tend to use heroes as tools and means to publicize and spread socialist core values. In dealing with the relationship between the government and the people, attention should be paid to two aspects. On the one hand, the concept of folk heroes

has accumulated profound genes of China's excellent traditional culture, so we need to creatively transform and innovatively develop traditional Chinese hero culture based on new historical conditions; on the other hand, To provide efficient instrument to resist the heroes of emperors, generals, and ministers, as well as the heroes of divine and martial arts revival, it is essential to shape heroic figures that embody the core values of socialism—scientific and cultural heroes, democratic and law-abiding heroes, and heroes of fairness and justice.

5 The Concept of Heroes in China in the New Era Takes Shape Within the Evolving Framework of Modern Democracy and the Rule of Law

How should the concept of heroes, as guided by the dominant national ideology in the new era, align with modern democracy and the rule of law? How should the construction of heroism be addressed within the framework of contemporary legal and democratic institutions? These questions require attention to two key aspects.

Firstly, we need to pay attention to the modern transformation of China's traditional conceptions of Heroes. After reform and opening up, China is on the road to full modernization, in which the modern transformation of the hero view is inevitable. The modern legal system has established such rules as that natural persons and non-law-enforcement agencies have no right to enforce the law, and that excessive defense is also a crime, making it possible to help others in the traditional sense of the word to commit a crime, and bringing new issues to the concept of the hero in the new era. Modern heroes differed from pre-modern heroes in that the former were heroes who were regulated and recognized by law, not heroes who broke the law, while the latter, on the contrary, were individuals who acted against the spirit of the law in the absence of the rule of law, on the basis of moral righteousness and in disregard of the right to life of others. If the study of heroism is detached from the context of modern legal frameworks, it risks falling into a predicament that contradicts the spirit of contemporary law. The widely popular CCTV drama In the Name of the People and films such as The Story of Qiu Ju, I Am Not Madame Bovary, and Mr. Six have vividly brought the issue of "rights protection" to the forefront for audiences. Therefore, examining the types, mechanisms of formation, and functional characteristics of modern heroism from a legal perspective has become a matter of critical importance.

Secondly, to ensure that heroic actions are not trapped in moral dilemmas or legal predicaments, the state must provide legal protection and safeguards for heroes at the institutional level. Numerous cases have shown the precarious situation of individuals who demonstrate bravery in defending others, only to be wrongfully accused or legally disadvantaged. Such tragedies expose gaps in the legal and social security systems, necessitating stronger legislative measures to protect heroes from unjust repercussions. The enactment of civil laws in 2017 to protect the reputations of heroes underscores the necessity of incorporating legal enforcement mechanisms into the theoretical construction of heroism.

President Xi Jinping has repeatedly emphasized the importance of honoring and protecting heroes. During his address at the plenary meeting of the delegation of the

People's Liberation Army and the People's Armed Police Force at the 13th National People's Congress, he stated: "Soldiers are the most lovable people. Ensuring that they receive the respect they deserve is a fundamental necessity. We must effectively manage and protect retired military personnel, ensuring that their rights are upheld and that relevant policies are fully implemented. Heroes must not be left to shed both blood and tears." The promulgation of the Law of the People's Republic of China on the Protection of Heroes and Martyrs institutionalized legal safeguards for national heroes, serving as a strong response to historical nihilism, the deconstruction of classical narratives, and the defamation of heroic figures.

Similarly, the Outline for the Implementation of Citizen Moral Development in the New Era, issued by the Publicity Department of the Communist Party of China on June 25, 2019, called for the "respect, praise, and care for outstanding figures and heroic models, as well as the preservation of their honor and image, thereby fostering a value orientation where virtue is rewarded." [9] In response to the increasing phenomenon of hero-bashing and historical nihilism, Xi Jinping has consistently stressed the need to revere and defend heroes, declaring: "We must never engage in acts that desecrate our ancestors, distort classical narratives, or defame our heroes." He further stated: "We must remember all those heroes who have contributed to the Chinese nation and the Chinese people. We must honor, defend, study, and care for our heroes. Together, with concerted efforts, we shall strive toward the realization of the 'Two Centenary Goals' and the great rejuvenation of the Chinese nation." [1] In various public engagements, including meetings with veterans, commendation ceremonies for model heroes, and memorial events for martyrs, Xi has underscored the necessity of learning from and emulating heroic figures: "We must remain steadfast like heroes and strive to contribute as they did." [10].

6 The Concept of Heroes in China in the New Era is Continuously Shaped Through Literary Creation and Criticism

The construction and dissemination of the concept of heroes under the dominant national mainstream ideology in the new era should emphasize the concept of heroes in literary and artistic creation and literary and artistic criticism, as reflected in the following two aspects.

On the one hand, the types of heroes in literary creations in the new era are increasing, and the space for heroes is expanding. The construction and dissemination of China's view of heroes in the new era should deal with the relationship between unity and diversity. Compared with the previous single combat heroes, revolutionary heroes and moral heroes, the new era has seen the emergence of "uderdog heroes", "commoner heroes", "tech heroes", "reform heroes", "capital heroes", "heroic monarch", "Sci-fi heroes «and other types, and the way of modeling has also changed. There are also new changes in the way they are portrayed.

On the other hand, as previously discussed, under the influence of postmodern trends that seek to deconstruct heroism, the aesthetic reconstruction of heroic figures in contemporary times must restore their sublime and transcendental essence. President Xi Jinping, in his speech at the Tenth National Congress of the China Federation of Literary and

Art Circles and the Ninth National Congress of the China Writers Association, emphasized: «For the heroes of the Chinese nation, we must hold deep reverence, depict and celebrate them with vivid and powerful expression, ensuring that heroes are honored in literary and artistic works. This will guide the people in establishing a correct historical perspective, national consciousness, patriotic sentiment, and cultural outlook. We must never desecrate our ancestors, distort classical narratives, or defame heroic figures." [11] At a symposium on literary and artistic work, he further stated: "Some works mock the sublime, distort classics, subvert history, and vilify the people and heroic figures." [12] The aesthetic essence of heroism lies in humanity's transcendental pursuit of the sublime. However, heroism is not merely a matter of pure aesthetics—its representation is inherently intertwined with desire, politics, and morality, forming three key dimensions of heroic expression. The genealogy of heroism has undergone a transformation from mythological manifestations to political figures, then to moral exemplars, and finally to entertainment-oriented depictions, resulting in various modes of heroic representation. However, in the consumer-driven postmodern era, heroism has been deconstructed and trivialized—its fusion of divinity and humanity, its political idealism, and its moral praxis have been dismantled. The transformation of the sublime into the grotesque has stripped heroism of its grandeur, leaving only a hollowed-out caricature of mockery and aesthetic degradation.

Postmodern deconstruction has reduced heroic figures to parodic substitutes, replacing the once grand and idealized heroic archetypes with rogues, prostitutes, the disabled, jesters, and cooks. Simultaneously, the postmodern practice of "de-heroization" not only diminishes the stature of heroes but also erases their moral integrity [13]. As the deconstruction and parody of heroism have become prevailing cultural trends, it is imperative to restore the sublime and transcendental essence of heroic figures in contemporary aesthetic reconstruction.

At a time when heroes are being portrayed all over the world, commentary on heroes is in a state of decline. Especially in the virtual space, animation images, online games and other projects that the majority of young people like and participate in, the synchronization of the heroes of the commentary is almost zero, the need for heroes of the commentary to take on the criticism of the wrong ideas, the "anti-hero" argument to clean up the phenomenon of the rise of the phenomenon. The more people speak out in society, the more a voice is needed to guide the public to form a correct view of heroes. It is necessary to study how to use the new media as a carrier for the dissemination of the concept of heroes, as the main channel for promoting the correct concept of heroes, and to deliver timely, fresh and valuable comments on heroes through modern means of communication. Combine this with new forms of communication and innovations in communication technology in order to attract more people to understand and embrace the socialist core values embedded behind heroes.

7 The Systematic Construction of the Concept of Heroes Under the Dominant National Ideology in the New Era

In the discussion above, we have examined the construction and dissemination of the concept of heroes in the new era under the dominant national ideology through an integrated, multidimensional research perspective that interconnects ideas and communication, the

global and the national, the official and the popular, as well as literature and philosophy. Based on this framework, we now turn to the foundational rationale, dissemination mechanisms, structural dimensions, and developmental goals of the concept of heroes in China in a new era.

The "foundational rationale" of the concept of heroes in China in a new era stems from the nation's practical need for heroes in its developmental process. The new era marks a new stage in the advancement of Socialism with Chinese Characteristics. President Xi Jinping has stated: «Today, China is undergoing rapid and unprecedented changes. We are closer than ever in history to realizing the great rejuvenation of the Chinese nation. To achieve our goals, we need heroes, and we need the spirit of heroism." [1] As a symbolic representation of China's core values, the promotion, study, defense, and care of heroic figures constitute a fundamental national policy aligned with the country's practical needs and developmental trajectory. The new era is both the social reality we currently inhabit and the starting point for a pragmatic and empirical approach to thought. Since the reform and opening-up period, with the transformation of China's market economy system, the country has developed a "new era" social structure that encompasses pre-modern, modern, and postmodern elements in a simultaneous coexistence. In this historical process, Western modernist and postmodernist values—including the concept of heroes—have also entered China and exerted an influence, giving rise to a pluralistic heroic space. Thus, the fundamental challenge in constructing and disseminating the concept of heroism in China in the new era lies in balancing the relationship between "unity and diversity."

The "dissemination mechanisms" of the concept of heroes in China in a new era must be rooted in modernized, all-media communication strategies. The application of communication theories should be leveraged to refine and optimize the construction and dissemination of the concept of heroism, effectively enhancing its communicative impact under the guidance of the dominant national ideology. Instead of relying solely on traditional propaganda-oriented approaches, the study of heroism's dissemination should be conducted from a comprehensive, media-centered perspective. Moreover, considering the contemporary globalized, digitalized, and media-integrated socio-cultural landscape, efforts must be made to enhance and expand the effective dissemination models of heroism through a multidimensional, interdisciplinary, and in-depth approach.

The "structural dimensions" of the concept of heroes in China in a new era should encompass multiple, integrated perspectives, including sociological, communicational, global, modernity, legal, philosophical, and literary dimensions. By incorporating these various aspects, a theory of heroism can be established that is not only distinctively characteristic of the new era in China but also capable of guiding the dissemination of the concept of heroism in contemporary society. The construction of this heroism theory must demonstrate originality, national specificity, adaptability, and global relevance.

The direction and objectives of the concept of heroes in China in a new era are reflected in the following aspects:

Firstly, at the domestic level, the concept of heroism in the new era should be firmly grounded in the core values of socialism, guiding social development and inspiring the people to move forward.

Secondly, in terms of international discourse, the concept of heroism in the new era should be established on the foundation of a community with a shared future for mankind, connecting the national and the global, and integrating with the universal spirit of heroism. Through comparative studies of different global hero archetypes, the goal is to identify a universally communicable core value system based on fairness and justice, fostering a global consensus that "the people are the heroes" and consolidating a collective commitment to fairness and justice.

Finally, in both constructing a socialist core value-oriented heroism and shaping a heroism that embodies the responsibilities of a community with a shared future for mankind, as well as in the organic unification of these two aspects, it is essential to develop a new theoretical framework that prominently highlights scientific and cultural heroes, democratic and law-abiding heroes, and heroes of fairness and justice—ordinary heroes as both moral exemplars and artistic archetypes. This framework will define the new goals and directions in the theoretical construction of heroes in a new era.

As Socialism with Chinese Characteristics enters a new era, President Xi Jinping put forward the important assertion that "the people are the creators of history, and the people are the true heroes." [14] He integrated the concepts of heroes and ordinariness, emphasizing that "greatness is forged in the ordinary, and heroes emerge from the people." [15] This laid the foundation for the "People's Heroism" under the guidance of the dominant national ideology in a new era. President Xi Jinping stated: «Greatness comes from the ordinary, and heroes emerge from the people. Completing every ordinary task with dedication is extraordinary in itself. The great cause of Socialism with Chinese Characteristics in the new era requires countless groups of heroes and heroic individuals. Learning from heroic deeds and carrying forward the spirit of heroism means embodying the extraordinary spirit of heroism in ordinary workplaces. «In his 2024 New Year's address, President Xi Jinping highlighted: "Diligent farmers, hardworking workers, bold and enterprising entrepreneurs, soldiers safeguarding the nation—people from all walks of life are striving with dedication. Every ordinary person has made an extraordinary contribution! The people are always our greatest source of strength in overcoming all difficulties and challenges." "Countless ordinary heroes are fighting and striving, converging into the mighty current of China's progress in a new era." [16].

In summary, starting from the new era's societal demand for heroes in China, we have explored the fundamental principles of the relationship between heroism and communication, the integration of national heroism with global heroism, the interaction between official and popular conceptions of heroism, the connection between modern democratic and legal heroism and traditional heroism, as well as the intrinsic nature and patterns of heroism in literary creation and criticism. Our aim is to examine the ontological foundations and preconditions for the construction and dissemination of heroism under the guidance of the dominant national ideology. We seek to develop a theory of heroism that is not only distinctly characteristic of the new era in China but also capable of guiding the dissemination of heroism in contemporary society. This theoretical framework should embody originality, national specificity, adaptability, and global relevance, thereby establishing a uniquely Chinese highly effective model for the promotion and communication of heroism in the new era.

Funded Project. National Social Science Fund General Project "Research on the Changes in Chinese Social Forms and the Evolution of Sublime Aesthetic Forms Since Modern Times" (21BZW188).

References

1. Xi, J.: Xi Jinping's Speech at the Series of Activities Commemorating the 70th Anniversary of the Victory of the Chinese People's War of Resistance Against Japanese Aggression and the World Anti-Fascist War, p. 19. People's Publishing House (2015)
2. Xi, J.: Secure a Decisive Victory in Building a Moderately Prosperous Society in All Respects and Strive for the Great Success of Socialism with Chinese Characteristics for a New Era: Delivered at the 19th National Congress of the Communist Party of China, p. 10. People's Publishing House (2017)
3. Wang, J.: What happens when bie-morden encounters space and time. J. CUPL **32**(3), 32 (2018)
4. Wang, J.: The heroic space and heroic rebirth in the postmodern context. Lit. Rev. **3**, 159 (2014)
5. Laozi: New Commentary and Translation, Translated and Annotated by Chen Guying, p. 226. Commercial Press (2004)
6. Lance, S., Hu, J.: Heroes and/as communication. J. Shanghai Univ. (Soc. Sci. Edn.) (3), 119 (2016)
7. Central Compilation and Translation Bureau (ed.): A Compilation of Key Articles and Documents Since the 18th National Congress of the Communist Party of China, vol. 2, p. 205. Central Party Literature Press (2016)
8. Xi, J.: Speech at the Commemoration of the 80th Anniversary of the Red Army's Long March, p. 2. People's Daily (2016)
9. Central Committee of the Communist Party of China and the State Council: Outline for the Implementation of Civic Moral Construction in the New Era, p. 1. People's Daily (2019)
10. Xi, J.: Speech at the Ceremony of Conferment of the National Medals and Honorary Titles, p. 2. People's Daily (2019)
11. Xi, J.: Speech at the Opening Ceremony of the 10th National Congress of the Chinese Federation of Literary and Art Circles and the 9th National Congress of the Chinese Writers' Association, p. 2. People's Daily (2016)
12. Xi, J.: Speech at the Symposium on Literary and Artistic Work, p. 2. People's Daily (2014)
13. Xu, D.: Heroes and the sublime as aesthetic forms. Guizhou Soc. Sci. **1**, 32–37 (2017)
14. Xi, J.: Speech at the 'Do Not Forget the Original Intention, Keep in Mind the Mission' Themed Education Work Conference, p. 4. People's Publishing House (2019)
15. Xi, J.: New Year's Speech by Xi Jinping, President of the People's Republic of China, p. 1. People's Daily (2021)
16. Xi, J.: New Year's Speech by Xi Jinping, President of the People's Republic of China, p. 1. People's Daily (2024)

A Study on the Installation Design Path of Song Dynasty Aesthetics in Narrative Exhibition Space

Yani Yang, Zhiyuan Lu, and Qi An(✉)

Beijing City University, Beijing 101309, China
anqi@bcu.edu.cn

Abstract. Exhibition space is crucial for spreading design ideas and installation design practice. Narrative space design can improve user experience. The Song Dynasty culture has a far - reaching influence. By integrating Song - style aesthetic elements, VR technology, and engaging visitors' five senses, we can better evoke their emotional resonance. From the Song Dynasty aesthetics view, this paper explores the changes in visitors' emotions and experiences in exhibition - space installation art design. It also studies the combination with modern technology to boost the narrative and appeal of historical spaces. Using literature analysis, case analysis, interdisciplinary methods, and inductive analysis, it aims to figure out how to promote visitors' high - level participation and emotional interaction in installation - art visits. It blends Chinese culture and aesthetics from the Song Dynasty's artistic charm into exhibition - space design. The installation design of narrative exhibition spaces with Song - Dynasty - aesthetic charm is significant in ideological basis, spatial display, and practical value. Narrative techniques can create scenes, enhance user experience, and build a deep - level interactive system for installation design.

Keywords: Narrative Space · Installation Design · Song Dynasty Aesthetics · User Experience · Design Philosophy

1 Background

In the current digital and information - age context, exhibition space is not only a platform for information dissemination but also an important carrier for emotional communication and cultural inheritance. With the continuous improvement of people's aesthetic demands, traditional exhibition methods can hardly meet the modern audience's pursuit of immersive experiences. Therefore, exploring innovative design concepts and methods has become the key to enhancing the attractiveness of exhibition spaces. Song Dynasty aesthetics is an artistic field worthy of in - depth study. With its unique artistic charm and aesthetic ideas, it provides new sources of inspiration for the installation art design in modern exhibition spaces. This research aims to explore the application path of Song Dynasty aesthetics in the installation design of narrative exhibition spaces by integrating the synergistic relationship among the three elements of people, space, and culture. By

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 210–219, 2025.
https://doi.org/10.1007/978-3-031-94171-9_18

incorporating the characteristics of Song Dynasty aesthetics, such as simplicity, naturalness, and profound artistic conception, into modern exhibition space design, it is expected to achieve emotional, interactive, and story - telling spaces, enhance the user experience, increase visitors' cultural identity, and promote the diversified development of installation design in exhibition spaces.

2 Literature Review

Installation art, as a modern art form that emerged in the mid - 20th century, has a wide - ranging and unique connotation [1]. Generally speaking, installation art refers to the act that artists endow objects placed, transformed, or fabricated in an environmental space with themes and ideas, and display the intended effects through light, shadow, sound, smell, materials, etc., thus realizing its value and concept [2]. Transcending the visual limitations of traditional art, it uses space, time, and the environment as carriers and skillfully integrates multiple perceptual elements, aiming to touch the emotions of visitors and guide and inspire their thinking.

At present, most interactive installation design cases mainly focus on enhancing the sense of technology and artistry, but they are rather weak in triggering emotional resonance among users. The development direction of installation art design should not merely focus on the accumulation and display of technologies. Instead, more attention should be paid to how to establish an emotional connection with visitors [3], better integrate traditional culture with modern technology, create exhibition spaces that are both culturally profound and contemporary, and enhance visitors' sense of identity.

The installation design in exhibition halls pays more attention to the emotional experience of visitors themselves. On this basis, adding the artistic conception of Song Dynasty aesthetics and the support of narrative theory can effectively increase visitors' participation and emotional resonance.

In many design fields, cultural aspects are taken into consideration. Only by centering on users and taking behavioral culture as the axis can we understand the spiritual cultural level above and the material cultural level below. As shown in the figure, based on the relational framework of "User - Space - Culture", in order to promote the interactivity between visitors and cultural spaces and achieve the goal of enhancing identity, the research on cultural - themed installation design is an indispensable part at this stage (See Fig. 1).

Narrative refers to the act of recounting stories and plots, and it conveys specific events that creators want to express through language texts or other forms. Space, on the other hand, is the area where humans live and carry out various activities. It is an objective entity and is centered around humans. The choice of narrative perspective directly affects the credibility of the narrative, the transmission of information, and the emotional engagement of the audience [4]. In fact, it is also the product of the integration, reconstruction, and organic combination of multiple media types and different spatial levels. A complete narrative space requires the joint efforts of creators and visitors. The combination of time and space plays a significant role in narration. Designers need to analyze the touchpoints of real audiences during the experience of narrative situations from the dimensions of time and space [5]. The integration of time and space promotes

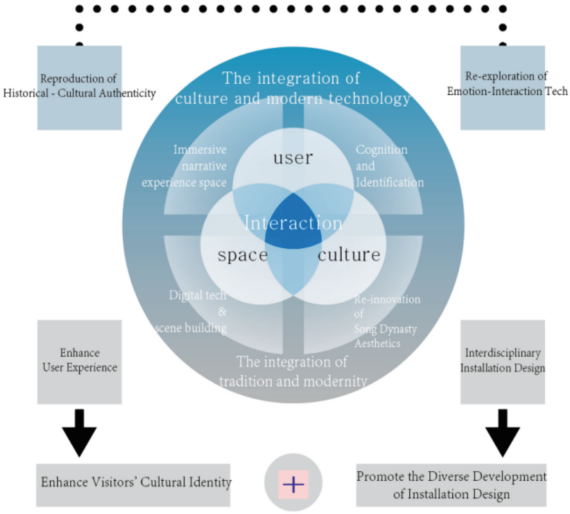


Fig. 1. A figure caption is always placed below the illustration. Short captions are centered, while long ones are justified. The macro button chooses the correct format automatically.

the progress of narration, abandoning the previous time - only narrative. The space in narration is not only visible but also a perceptual space for visitors, an immersive space that can be experienced by visitors and trigger emotional resonance. The core of the immersive experience in a narrative space is "experience," and "immersive" is its presentation state. For an immersive experience, the focus is on sensory experience and interactive experience. The sensory experience mainly relies on the stimulation of the five senses - vision, hearing, smell, taste, and touch - to completely "envelop" the audience and immerse them in it. To create this deep sense of immersion, the audience's five senses need to perceive the digital environment as real as the physical reality. At present, the installation design in exhibition spaces usually integrates the "five - sense" experience and adds somatosensory elements [6]. There is no doubt that the multi - sensory immersive experience of creating scenarios and narrative stories in the entire exhibition successfully enhances the connection with visitors. The emergence of various new technologies has had a profound impact on the development of the display forms and interactive forms of immersive experiences [7].

The aesthetic elements of the Song Dynasty profoundly influence modern design. Many design fields blend Song Dynasty aesthetics with modern design, creating works that merge modern aesthetics and cultural connotations. Understanding the internal logic and artistic expression of Song Dynasty aesthetic thoughts guides the application of aesthetic ideas and the development of space design. Research on color and space in Song Dynasty aesthetics has a strong emotional and suggestive impact, presenting cultural ideas, with strong symbolism, and making history more multi - dimensional. It's an important extension of historical iconography.

This study integrates the core concepts and spiritual culture of Song Dynasty aesthetics, analyzes its uniqueness in shape treatment, material expression, and color application, explores the combination of these aesthetic elements with narrative spaces, and designs installation artworks in exhibition spaces. These artworks enhance visitors' participation and emotional resonance, with a greater emphasis on users' experiences in the installation design.

3 Methods

This research uses literature analysis, case analysis, interdisciplinary and inductive methods.

For key terms like "narrative space", "Song Dynasty aesthetics", "exhibition space" and "installation design", relevant literature is searched, sorted and analyzed. By understanding research status and results, research goals are set, laying a theoretical groundwork and defining the research's value.

Song - Dynasty - aesthetics - themed design cases are analyzed, focusing on design methods and concepts. Data is explored to summarize common scheme - design methods from the Song - Dynasty - aesthetics and narrative - space - installation - design perspectives.

Based on narrative - space theory, user - experience - based spatial design is applied, integrating disciplines like communication and narratology. Design features related to Song Dynasty aesthetics are analyzed to support narrative - installation design. Interactive - technology methods are also covered for technical support. Cross - disciplinary learning helps apply new ideas to research for practical verification.

User interactive behaviors are observed, analyzed inductively, and investigation results are summarized to get design elements, principles and strategies.

4 Discussion

4.1 The Application of Interaction in Installation Design

The concept of human - computer interaction is widely applied in installation design. There are already many successful cases of interactive installation design, which create different interactive installation designs tailored to different groups of people and their needs (As shown in Table 1).

4.2 Extraction of Design Elements

Song Dynasty aesthetics has distinct characteristics of its time. Its forms of expression are analyzed from three aspects: shape treatment, material expression, and color extraction.

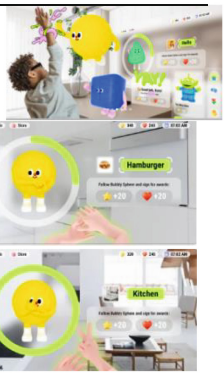
In terms of shape treatment, the traditional architectural language of Song Dynasty aesthetics undergoes modern deconstruction. The specific design manifestations include "points, lines, and planes", blank spaces with modern geometric elements, and the undulation of points on multiple spatial interfaces [8].

Under the influence of Song Dynasty aesthetics, the application of design materials and colors is unique and contains a profound cultural core.

In the design of Song Dynasty aesthetic elements, these key terms become the core of constructing the spatial framework.

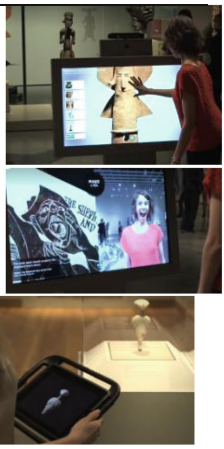
From the perspective of materials, wooden and bamboo materials are natural and simple, with clear textures, conforming to the simple lifestyle concept of scholars in the Song Dynasty. Their warm and smooth textures bring a soft touch to the space. Stone bricks are solid and stable. When used in architectural sculptures, they highlight the grandeur of Song Dynasty architecture. The ceramic craftsmanship was highly developed in the Song Dynasty. For example, the sky - blue porcelain of Ru Kiln is delicate and colorful, showing the exquisiteness of the material and reflecting the Song Dynasty's pursuit of a refined life (as shown in Table 2).

Table 1. Human - Computer Interaction Installation Design.

Design Name	Design Introduction	Design Configuration	Design Features and Content	Design
SignyFriendy	An interactive system for autistic children enables them to enter SignyFriendy's world via AR glasses and smart gloves. This system also helps unlock their learning potential through sign language exploration.	AR glasses, smart gloves, virtual IP characters	Visual and auditory sensory interactive experiences create an immersive space that mobilizes the emotions and physical rhythms of users.	
LUMES Installation	LUMES is an interactive LED digital wall art at the entrance of Cabrini Hospital's pediatric ward. It's a luminous wall interactive system, integrating with the surrounding architecture and featuring clear colors.	Motion capture, LED digital wall	Tactile and visual sensory interactions, with dreamy graphics, delight children, patients, and staff.	

(*continued*)

Table 1. (*continued*)

Design Name	Design Introduction	Design Configuration	Design Features and Content	Design
Design of Digital Exhibits and Art Installations in Cleveland Museum of Art - Gallery One	In this groundbreaking exhibition, Gallery One integrates art, technology, and interpretation, inspiring visitors to explore the museum's permanent collections.	Digital interactive touch screen Gesture - mimicking interaction Virtual canvas experience Dynamic capture creation	Interactive installation design enables interactions between visitors and museum staff, between visitors and museum culture, and among visitors. This deepens visitors' impression of exhibits and their cultural identification.	

In the Song Dynasty, soft, elegant colors were favored over bright ones. Light, subdued hues like ochre, azurite, malachite green, and indigo blue were used to evoke a subtle, profound atmosphere (As shown in Table 3).

In terms of modeling, taking natural landscapes as the theme, with large - scale blank spaces, it showcases the pure beauty of nature through smooth lines and harmonious proportions. The garden architecture borrows scenery to conform to nature, creating a profound artistic conception. The traditional architectural language of Song Dynasty aesthetics undergoes modern deconstructive modeling treatment. The specific design manifestations include “points, lines, and planes”, blank spaces with modern geometric elements, and the undulation of points on multiple spatial interfaces [9].

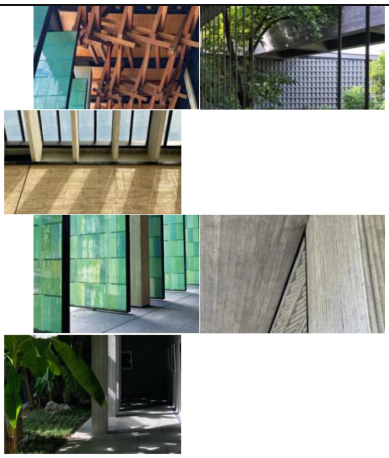
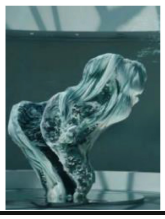
These elements jointly construct the spatial framework of Song Dynasty aesthetics, making people seem to be in the time - space of the Song Dynasty and feel its unique charm.

4.3 Design Strategies

Scene Creation to Stimulate Sensory Experiences. Driving visitors’ visits mainly through scenarios and narrative spaces means a more profound immersive experience. Virtual technology has made immersion more common and accessible in exhibition spaces, and soon the word “immersion” has been used more frequently in exhibition spaces. Visitors in the 21st century prefer immersive experiences rather than passive observation or acceptance.

In physical interactions, a more natural operation method can be attempted [10].

Table 2. Material Extraction of Designs with Aesthetic Elements of the Song Dynasty.

Design Works	Work Titles	Material Application
	Hangzhou branch of the National Archives of Publications and Culture	Timber structure, brick, rammed earth, celadon, bamboo - formwork concrete, Chinese plants
	"The Celestial Blue Tranquility" Installation Design	Simulate the texture of Ru kiln porcelain

By applying Song Dynasty aesthetic ideas in installation design, these ideas are fully presented via elements like material textures and colors. This stimulates visitors' five senses, creating a unique scene of that era. Mainly natural materials are chosen. After processing, the natural beauty of decorative shapes is accentuated. Materials such as air - dried reeds and raw wood prompt visitors to touch them, strengthening their memories of Song - Dynasty - themed spaces.

Meanwhile, using simple and intuitive materials and forms, a simple, reserved, and subtle atmosphere is crafted within the installation space. This showcases the "taste of plainness" in Song Dynasty aesthetics, achieving a return to nature, conforming to natural laws, highlighting artistic charm, and forming a man - nature - unity aesthetic conception.

The use of unique Song Dynasty aesthetic colors is a crucial part of creating the atmosphere. We should give full play to our initiative, follow the spirit of Song Dynasty aesthetics, adjust spatial colors, use colors to give visitors psychological cues, express emotions with different colors, and create a harmonious color environment. Abundant use of natural colors can enhance the atmosphere and convey emotions, realizing the interaction between personal feelings and the environmental ambiance. When selecting

Table 3. Color Extraction of Paintings in the Song Dynasty.

composition	Color Name	color	RGB values	hexadecimal color code
A Thousand Miles of Rivers and Mountains (partial)	mineral blue		30,127,160	#1e7fa0
	mineral green		32,104,100	#206864
	ochre pigment		132,90,51	#845A33
	indigo		26,40,71	#1A2847
The Painting of Listening to the Qin(partial)	ochre pigment		132,90,51	#845A33
	mineral green		32,104,100	#206864
	jet black		0,0,0	#12090c
	cochineal red		158,43,46	#9E2B2E
Drawing of Rare Birds from Life	ochre pigment		132,90,51	#845A33
	mineral green		32,104,100	#206864
	indigo		26,40,71	#1A2847
	vermilion		233,72,41	#E94829
	wisteria yellow		255,182,30	#FFB61E

material colors, brown and tan are primary. These colors blend well with the natural environment, exude a natural flavor, and offer visitors an immersive visual experience.

Technological Interaction to Stimulate Entertainment Experiences. With the rapid development of computer technology, exhibition spaces have broken through the limitations of space and time. Installation design can connect the real world and the virtual world simultaneously, with more powerful expressiveness.

The net artist Wang Bo analyzed "interaction". He believes that "structural interaction" exists with interactivity as the main structure of the work. It cannot be directly seen by users but can be deliberately designed and hidden in the work by creators. It can reflect the creators' intentions to the greatest extent [11]. People should be regarded as the initiators of the entire process [12], which means that visitors should be given primary consideration. By means of VR, AR, and 3D online exhibition halls, the artistic and authentic display and narration of the Song Dynasty are carried out, making visitors participants in the narrative space and having a substantial interaction with the installation design. Visitors scan the QR code at the entrance of the exhibition hall, and their mobile phones provide a camera interface. After entering the installation space, when they use the camera interface of their mobile phones to photograph the installation design, scenes with Song Dynasty aesthetic elements will appear, and they can further design the scenes to increase their participation (See Fig. 2). For installation design, using scene restoration to connect objects, people, and spaces within the space, from a practical perspective, visitors can transcend the limitations of time and space and obtain a unique cultural experience.

Fig. 2. The Visiting Process and Interactive Behaviors of Visitors.

5 Conclusion

This research explores Song Dynasty aesthetics in narrative exhibition space installations, showing its potential to blend with modern design. Using the simplicity and naturalness of Song aesthetics can boost the space's aesthetic, resonate with the audience, and add charm. It also highlights VR and AR's key role in creating immersive interactions. The study of Song - era color and space have emotional and symbolic meaning, extending historical presentation.

In the digital era, exhibition space installation design has opportunities and challenges. Integrating tech with aesthetics to create culturally - rich, contemporary spaces for better user experience is a new design topic. This research offers an innovation reference, expecting more creative and meaningful exhibition space designs.

Acknowledgements. Research funded by the University-Industry Collaborative Education Program of Beijing City University (CXHZ2025001)

References

1. Li, W.: Space re-creation: the integration and transformation of installation art and museum exhibition design. Southeast Cult. (S1), 42–47 (2023)
2. Wang, R., Wang, X.: The application of installation art in exhibition space design. Ind. Des. (03), 91–93 (2023)
3. Ma, L.: The application of interactive design in museum exhibition spaces in the user experience era. Beauty Times (Upper) (08), 32–35 (2023)
4. Genette, G.: Narrative Discourse: An Essay in Method. Cornell University Press, New York (1980)
5. Chien, C.-F., Kerh, R., Lin, K.-Y., Yu, A.P.-I.: Data-driven innovation to capture user-experience product design: an empirical study for notebook visual aesthetics design. Comput. Ind. Eng. **7**(6), 162–173 (2016)
6. Classen, C.: The Deepest Sense: A Cultural History of Touch. University of Illinois Press, Champaign (2012)
7. Scholz, J., Smith, A.N.: Augmented reality: designing immersive experiences that maximize consumer engagement. Bus. Horiz. **59**(2), 149–161 (2016)
8. Li, Y.: Design of B&B with song-style aesthetics. Constr. Econ. **44**(08), 106 (2023)
9. Wang, Y.: An analysis of the influence of song-style aesthetic colors on modern home furnishings. Zhejiang Arts Crafts **02**, 100–102 (2024)
10. Kroll, H., Pirklbauer, J., Kalo, J.C., et al.: Narrative query graphs for entity-interaction-aware document retrieval. In: International Conference on Asian Digital Libraries, pp. 80–95. Springer, Cham (2021)
11. Bourdot, P., Convard, T., Picon, F., et al.: VR–CAD integration: multimodal immersive interaction and advanced haptic paradigms for implicit edition of CAD models. Comput.-Aided Des. **42**(5), 445–461 (2010)
12. Pisciella, S.: Peter Eisenman Gher-Ghar: House 11a. Mimesis, Milano (2018)

Interacting with GenAI and LLMs

Bridging UI Design and Chatbot Interactions: Applying Form-Based Principles to Conversational Agents

Sanjay Krishna Anbalagan[1(✉)], Xinrui Nei[2], Umesh Mohan[1], Vijay Kumar Kanamarlapudi[1], Anughna Kommalapati[1], and Xiaodan Zhao[2]

[1] Amazon Web Services, Dallas, USA
ameejay@amazon.com
[2] Amazon Web Services, Seattle, USA

Abstract. Domain-specific chatbot applications often involve multi-step interactions, such as refining search filters, selecting multiple items, or performing comparisons. Traditional graphical user interfaces (GUIs) handle these workflows by providing explicit "Submit" (commit data) and "Reset" (discard data) actions, allowing back-end systems to track user intent unambiguously. In contrast, conversational agents rely on subtle language cues, which can lead to confusion and incomplete context management. This paper proposes modeling these GUI-inspired metaphors—acknowledgment (submit-like) and context switching (reset-like)—as explicit tasks within large language model (LLM) prompts. By capturing user acknowledgment, reset actions, and chain-of-thought (CoT) reasoning as structured session data, we preserve clarity, reduce user confusion, and align domain-specific chatbot interactions with back-end logic. We demonstrate our approach in hotel booking and customer management scenarios, highlighting improvements in multi-turn task coherence, user satisfaction, and efficiency.

Keywords: GUI Inspired CoT · Submit|Rest metaphor · Domain-specific chatbots

1 Introduction

Q1: Why compare graphical user interface (GUI) interactions with chatbots?
A1: Traditional GUIs often provide explicit buttons for Submit (committing data) and Reset (discarding data). These interactions are unambiguous, allowing the back-end to track user intent clearly. In contrast, chatbots rely heavily on natural language cues that can be vague or underspecified, leading to confusion about user intent and context [1].

Q2: What challenges arise in domain-specific chatbot interactions?
A2: Domain-specific applications—like hotel booking systems, customer management tools, or inventory systems—frequently require multi-step user interactions. The chatbot must keep track of user selections (e.g., searching for a hotel, confirming a customer, refining filters). Without a clear "submit" or "reset" action, confusion can arise when

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 223–231, 2025.
https://doi.org/10.1007/978-3-031-94171-9_19

the user's language changes the context. For instance, a user might suddenly switch from discussing Customer A to Customer B without a clear prompt, leaving the system uncertain about which data to commit or discard [2].

Q3: How do Large Language Models (LLMs) come into play?
A3: LLMs can generate text (including structured data, such as XML tags) based on prompts. In traditional programming workflows, developers can prompt an LLM to produce specific variables or tags, parse these outputs, and then feed them back into the application's logic [3]. This creates a bridge between the flexibility of natural language and the clarity of explicit UI actions.

Q4: What solution does this paper propose?
A4: We propose modeling GUI-inspired metaphors—such as Submit and Reset—as explicit tasks within LLM prompts. By capturing user acknowledgment (Submit-like) or context switching (Reset-like) as structured session data, the system maintains clarity over whether the user wants to continue with the same context or switch to a new one. We also employ chain-of-thought (CoT) reasoning (see Q6) for multi-step tasks, enabling the LLM to clarify ambiguous details in user inputs and guide the back-end more transparently.

Q5: How does this approach benefit multi-turn chatbot interactions?
A5: By making "acknowledgment" and "reset" steps explicit (and incorporating step-by-step reasoning via CoT), the approach:

- Preserves clarity of context across many turns.
- Reduces user confusion and the potential for erroneous actions.
- Aligns domain-specific chatbot interactions with back-end logic.
- Improves coherence, user satisfaction, and efficiency in tasks like hotel booking or customer management.

Q6: How is chain-of-thought (CoT) integrated into your prompts?
A6: We instruct the LLM to produce structured intermediate steps, effectively revealing its internal reasoning about user context changes or form-like actions (e.g., "submit" vs. "reset" intentions). These steps are included in the LLM's output (potentially as JSON/XML segments or separate explanatory text) but are primarily used internally for clarity and debugging. By parsing this CoT output, the back-end can follow a transparent, step-by-step rationale: deciding whether to commit certain data, reset fields, or switch contexts based on the user's utterance. While the CoT can be hidden from end users, it allows developers to better understand how the LLM arrives at certain actions.

2 Related Work

Conversational interfaces have been extensively studied for tasks such as customer support, information retrieval, and personal assistance [1, 2]. Much of the prior work focuses on natural conversation flow or dialogue state tracking [4], which aims to maintain a belief state of user goals. However, these methods often do not explicitly treat commit (submit) or discard (reset) as separate metaphors—many rely on generalized user intent classification that can miss the nuance of partial context resets or confirmations central to domain-specific applications.

Recent advancements in Large Language Models have enabled systems to generate structured outputs that can be parsed in a programming context [3]. Efforts to unify LLM-based text generation with traditional application flow frequently highlight prompt design [4]—specifically, how to craft instructions so that the LLM's output is both semantically correct and programmatically useful. Our work builds on this line of research by introducing task-based prompting that mirrors GUI actions, augmented with chain-of-thought reasoning, thereby reducing ambiguity and preserving clarity in multi-turn dialogues.

3 Proposed Approach

3.1 Turning UI Actions into Prompted Tasks

Key Insight: In a typical GUI, Submit commits the data the user has entered, while Reset discards it, returning to a default or initial state. We replicate this idea in the chatbot context by designing specialized prompts that explicitly label user intentions. For example, when a user appears to confirm a specific customer, we interpret that as a Submit-like action. Conversely, when the user wants to abandon the current customer or switch to a new one, we treat it as a Reset-like action.

3.2 Task-Based Prompts (with CoT) for LLMs

LLMs can be prompted to produce structured data (e.g., XML or JSON) representing these actions. Our prompt engineering also incorporates chain-of-thought instructions, allowing the LLM to articulate intermediate reasoning steps (primarily for internal use). By parsing the model's output, we feed the resulting values (e.g., isCustomerConfirmed) and the CoT textual segments into the back-end. This ensures the system can reliably commit or reset user context with full transparency into the LLM's rationale.

Core Advantages

1. Unambiguous Context Management: Each prompt explicitly requests an action label (e.g., yes or no for confirming a customer context) and includes the intermediate reasoning steps to resolve ambiguities.
2. Seamless Integration: Developers can parse the LLM's output tags and the chain-of-thought textual segments (e.g., using a simple XML parser or regex) and integrate them into existing application logic.
3. Consistency Across Interactions: Multiple tasks (e.g., IsHotelSelectionConfirmed, IsBookingReset) can be introduced, each one governing a piece of the dialogue state, guided by CoT to clarify how each decision is reached.

4 Implementation

4.1 Example: Customer Confirmation Task

Below is an illustrative TypeScript code snippet that shows how we construct a Customer Confirmation Prompt. The goal is to detect whether the user wants to continue with the current customer or switch to a new one. We also provide "chain-of-thought" instructions so that the LLM's output includes a trace of its reasoning:

Human: You are a customer search bot, and your task is to determine if the user query refers to searching for a customer or details about the current customer based on <userQueryHistory> and <currentCustomerName>. Follow the guidelines:
1. Respond with <isCustomerConfirmed>no</isCustomerConfirmed> if:
- The user query mentions or implies the name of a different customer.
- The user query involves clarifying or correcting the current customer.
- The user query refers to a geographic or industry-specific refinement without explanation.
- The user query uses phrases like "the one", "I meant", or "I am looking for" that indicate a switch.
2. Respond with <isCustomerConfirmed>yes</isCustomerConfirmed> if:
- The user query asks about details of the current customer (service consumption, pricing, etc.).
- The user query does not involve clarifying or correcting the current customer name.
- The user query involves comparing details between the current customer and another but keeps the current customer context.
3. Include a brief explanation of your reasoning in <chainOfThought>...</chainOfThought> tags.
<examples>
<example>
query: "the one in china"
current customer name: ABCCompany
user query history list: ["Is ABCCompany a customer"]
<isCustomerConfirmed>no</isCustomerConfirmed>
<chainOfThought>User mentions 'the one in China', so likely switching context.</chainOfThought>
</example>
<example>
query: "recent news"
current customer name: ABCCompany
user query history list: ["Is ABCCompany a customer"]
<isCustomerConfirmed>yes</isCustomerConfirmed>
<chainOfThought>User is asking for details on ABCCompany with no mention of a new entity.</chainOfThought>
</example>
</examples>
Here is the user query:<query>${userQuestion}</query>
Here is the current customer name:<currentCustomerName>${customerName}</currentCustomerName>
Here is the user query history:
<userQueryHistory>
${JSON.stringify(userQueryHistoryList)}
</userQueryHistory>

How it works.

1. User Query: The user's most recent input.
2. Current Customer Name: The context the chatbot is currently using.

3. User Query History: A list of previous queries to help the LLM track context.
4. Response: The LLM returns both < isCustomerConfirmed > yes < /isCustomer-Confirmed > or < isCustomerConfirmed > no < /isCustomerConfirmed > and a < chainOfThought > … < /chainOfThought > block summarizing the reasoning. While the back-end can parse the chain-of-thought for logging or debugging, end-users typically only see the summarized result.

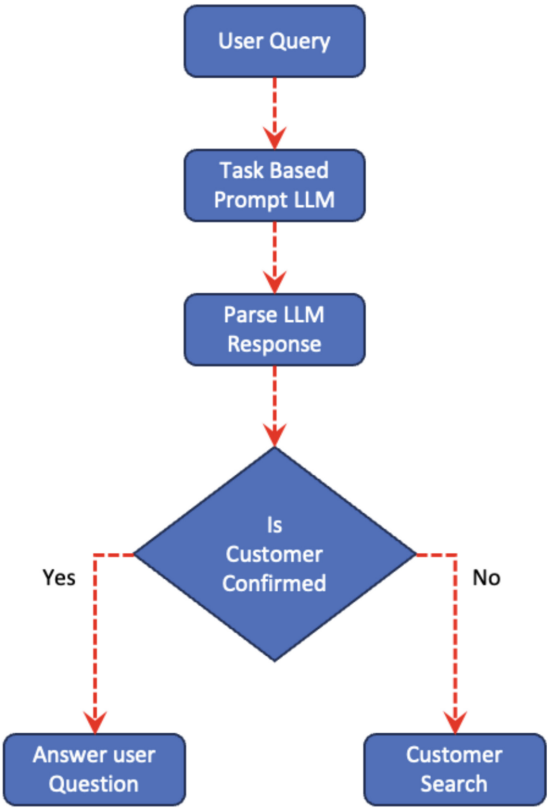


Fig. 1. Flow chart of the interaction.

Key Takeaway. The Explicit yes/no tag from the llm is interpreted much like a user clicking submit or reset in a graphical UI. The < chainOfThought > content provides a transparent, step-by-step rationale for each decision.

5 Example Scenarios

5.1 Example Conversation

Scenario. A user is searching for a specific customer

1. **User**: “Is ABCCompany a customer?”

LLM Output:

<isCustomerConfirmed>no</isCustomerConfirmed>
<chainOfThought>User is naming ABCCompany, no current customer context is set, so we treat it as a new search.</chainOfThought>

Action: The system searches for “ABCCompany”

2. **User**: “What’s their recent news?”

LLM Output

<isCustomerConfirmed>yes</isCustomerConfirmed>
<chainOfThought>User asks about 'their' recent news, referring to ABCCompany. Context remains ABCCompany.</chainOfThought>

Action: The system fetches the news for ABCCompany

3. **User**: “Actually show be XYZCompany info?”

LLM Output

<isCustomerConfirmed>no</isCustomerConfirmed>
<chainOfThought>User specifically wants XYZCompany info only, discarding ABCCompany context.</chainOfThought>'their' recent news, referring to ABCCompany. Context remains ABCCompany.</chainOfThought>

Action: The system resets the context to XYZCompany

5.2 Example Conversation with and with is Customer Confirmed (SUBMIT | RESET)

Without < isCustomerConfirmed > Tags

User: “Is Delta a customer?”
LLM: “Delta Airlines has been a customer for two years.”
User: “Dental.”
LLM: “Delta Airlines provides a dental plan for employees…”
User: “No, I meant Delta Dental.”
LLM: ”…Delta Dental might be a separate entity; let me see…”

Inference (Without)

No explicit "reset" cue is given. The system lumps "Dental" under Delta Airlines until the user explicitly clarifies "Delta Dental" as a new customer. This leads to confusion and extra steps.

With < isCustomerConfirmed > Tags

User: "Is Delta a customer?"
LLM :"Yes, Delta Airlines has been a customer for two years."
User: "Dental."
LLM: "Are you referring to Delta Dental as a new customer, or do you mean the dental plan from Delta Airlines?"
User: "I mean Delta Dental specifically."
LLM:"Understood. Let's talk about Delta Dental."

Inference (With)

Because < isCustomerConfirmed > evaluates to **no** on immediate follow up ("Dental"), the system realizes the user might be switching from "Delta Airlines" to "Delta Dental." It asks a clarifying question to confirm intent before pivoting the context. This ensures fewer misinterpretations and a smoother conversation flow.

6 Discussion and Future Work

6.1 Enhancing Multimodal Interactions

Although we focus on textual chat, the same approach can extend to voice-based interfaces or mixed interactions where users toggle between speech and minimal GUIs. Each "action" still needs to be recognized, tracked, and potentially explained in a chain-of-thought if the developer requires full transparency.

6.2 Generalizing to Other Domains

The approach generalizes beyond customer management to any domain with multi-step workflows—travel bookings, appointment scheduling, e-commerce carts. Whenever confirm (submit) vs. reset (cancel) is relevant, and step-by-step CoT clarifications are needed, this method can reduce ambiguity.

6.3 Limitation and Next Steps

LLM Reliability. The method depends heavily on the LLM consistently outputting correct tags and coherent chain-of-thought. prompt design and testing are crucial.

Edge Cases. Users might use ambiguous language or sarcasm, potentially causing faulty CoT or misclassification. fine-tuning and robust prompt engineering can mitigate this.

Preliminary Testing. In a small pilot study with 100 users, we observed ~ 30% fewer conversation misalignments (e.g., users having to restate or correct context) compared to a baseline chatbot approach. Future larger-scale tests could quantify improvements in user satisfaction and task completion time more rigorously.

7 Conclusion

We introduced a novel approach to modeling acknowledgment and context switching within domain-specific chatbot applications by repurposing GUI metaphors (Submit/Reset) as explicit tasks in LLM prompts. By capturing these actions as structured session data—such as < isCustomerConfirmed > yes < /isCustomerConfirmed > —and incorporating chain-of-thought reasoning (internally for clarity), we align the conversational flow with back-end logic more effectively. Our examples in customer management and hotel booking scenarios demonstrate how this mechanism helps mitigate user confusion, preserve coherent context, and create a more efficient, user-friendly interaction paradigm.

References

1. Radziwill, N., Benton, M.: Evaluating quality of chatbots and intelligent conversational agents. J. Quality Manag. **25**(2) (2017)
2. Nuruzzaman, M., et al.: A survey on chatbot implementation in customer service industry. Inter. J. Manag. Appli. Sci. (2018)
3. Williams, J. D., Zweig, G.: End-to-End LSTM-Based Dialog Control Optimized with Supervised and Reinforcement Learning. arXiv preprint (2016)
4. Brown, T., et al.: Language models are few-shot learners. Adv. Neural Inf. Process. Syst. (2020)
5. Reynolds, L., McDonell, K.: Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. arXiv preprint (2021)

Slow AI in Empowering and Understanding Marginalized Communities: Case Study in Campana-Altamira, Mexico

Savira Aristi[1(✉)], Hisa Martinez Nimi[2], Juan Carlos Chacon Quintero[2], Alejandro Martinez Baca[3], and Kenta Ono[2]

[1] Graduate School of Science and Engineering Chiba University, Chiba, Japan
23wd8306@student.gs.chiba-u.jp
[2] Chiba University, Chiba, Japan
[3] Tecnologico de Monterrey, Monterrey, Mexico

Abstract. This paper introduces "Slow AI," a human-centered, iterative methodology designed to integrate artificial intelligence (AI) into marginalized communities interventions sustainably. Unlike conventional AI systems, Slow AI emphasizes gradual learning, community participation, and socio-cultural sensitivity. By engaging directly with local residents and stakeholders, we explore how a deliberate and mindful integration of AI technologies can support sustainable development and address these pressing issues. The research explores experimental methodologies for data collection, modeling, and insight generation within the Campana-Altamira community in Monterrey, Mexico, a region facing significant socio-environmental challenges. Our approach combines the Locales Framework, Social Set Analysis, and Liquid Neural Network to ensure that information sharing and delivery formats are tailored to the community's social structures and specific contexts. The findings suggest that adopting a Slow AI framework fosters a deeper understanding of local contexts, promotes community empowerment, and mitigates the risks associated with rapid, one-size-fits-all AI deployments. This study also contributes to academic discourse by articulating the theoretical foundations of Slow AI, highlighting its potential to advance social equity and inclusion. While the framework is in its initial phase, it lays the groundwork for future iterations, emphasizing the importance of mutual learning and interdisciplinary collaboration in AI development for underserved communities.

Keywords: Slow AI · Marginalized Communities · Campana-Altamira · Data Empowerment · Intervention Design

1 Introduction: Slow AI

Slow Artificial Intelligence (AI) is an emerging paradigm that challenges the conventional emphasis on speed, efficiency, and large-scale automation in the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 232–242, 2025.
https://doi.org/10.1007/978-3-031-94171-9_20

development of AI. Drawing inspiration from movements like slow food and slow fashion, both advocate for sustainable practices that promote local communities, ethical production, and consumer consciousness [1–3], Slow AI advocates for a deliberate, human-centered approach to AI design and implementation, prioritizing community engagement, socio-cultural sensitivity, and ethical considerations [4].

Slow AI introduces the concept of 'slowness' to this fast-paced technology, understanding it beyond a temporal metric and rather as a change in modes of engagement that resist notions of speed, efficiency, and optimization rooted in its colonial and extractive history. It aims to challenge the prevailing trends of rapid consumption and the relentless pursuit of efficiency that have characterized the digital era, particularly within AI [5].

Incorporating insights from cognitive theories, some researchers propose AI architectures inspired by Daniel Kahneman's "Thinking, Fast and Slow," [6] integrating both rapid, experience-based responses and deliberate, analytical reasoning. This dual-process approach aims to enhance AI's adaptability, generalizability, and decision-making processes, aligning with the principles of Slow AI by allowing systems to evolve over time and gradually transition from slow to fast thinking with sufficient experience [7].

Recent research on AI highlighted the need for deliberative governance, where AI design choices are treated as political decisions reflecting societal values [8]. This aligns with the principles of transparency, accountability, and ethical responsibility in AI systems, particularly in light of their potential social impacts [9]. Although significant efforts have been made to address algorithmic biases, such as those led by organizations such as the Algorithmic Justice League [10], many AI interventions still operate within top-down, efficiency-driven paradigms that overlook the importance of social negotiation and contextual adaptation in marginalized communities.

The need for Slow AI emerges in response to these limitations, emphasizing deliberation, iterative refinement, and participatory governance in AI development. The approach of Slow AI aligns with calls for inclusive research practices [11,12] AI within existing knowledge ecosystems, allowing marginalized communities to shape, critique, and refine AI-driven insights in ways that reflect their lived experiences. Collaborative frameworks that integrate stakeholders, practitioners, and local communities [13,14] are critical in realizing this vision, as they shift AI from a deterministic tool to a deliberative partner in social impact research. Engaging communities in this process ensures that AI is not merely tailored to their needs but also aligned with their problem-solving practices, thereby counteracting systemic biases rather than exacerbating them [15–17].

The integration of AI in social impact research has been largely driven by efficiency, automation, and predictive accuracy. However, these conventional AI approaches often fail in complex, community-driven contexts, where knowledge is dynamic, trust is essential, and decision-making requires deliberation rather than immediate optimization. As data-driven systems become more embedded in decision-making, underserved communities often struggle with opaque AI models

that do not align with local knowledge, values, and problem-solving processes. Slow AI is an alternative paradigm that emphasizes deliberation, explainability, and participatory participation in speed and automation.

This research argues that slow AI is not a fixed technology or a specific tool, but rather a conceptual approach that can take multiple forms depending on the context of AI deployment. It is best understood as a design philosophy that seeks to slow down AI-driven decision making to allow reflection, negotiation, and social embedding. The specific operationalization of Slow AI depends on factors such as who the AI is meant to serve, the sociopolitical environment in which it operates, and the historical relationship between technology and trust within a given community. It is essential to recognize that a singular model of Slow AI may not be universally applicable to all marginalized communities, as these groups exhibit diverse characteristics and experiences that cannot be homogenized. The term "marginalized communities" itself has gained considerable traction in contemporary discourse, yet it encompasses a wide range of populations, each with unique socio-cultural contexts and challenges. The versatility of Slow AI as a meta-framework allows it to be tailored to specific technological, cultural, and social conditions, making it a valuable paradigm for ethical and community-driven AI development.

In marginalized settings such as Campana Altamira, collecting reliable and comprehensive data poses significant challenges due to limited access, socioeconomic barriers, trust issues, and infrastructural constraints [18]. In the Campana-Altamira case study, Slow AI is specifically implemented as an integral component of a broader participatory and iterative design process.

Our approach addresses key challenges in leveraging AI for social impact, especially in marginalized communities, specifically: (1) the lack of contextual awareness, (2) the overgeneralization of knowledge, and (3) the limited adaptability of AI to community-driven feedback. To address these limitations, we propose an integrative framework that combines Social Set Analysis (SSA) to model knowledge flows, the Locales Framework (LF) to ensure contextually relevant interactions, and Liquid Neural Networks (LNNs) to facilitate explainable and adaptive reasoning. This approach operationalizes Slow AI, emphasizing deliberation and iterative refinement through participatory feedback loops. By embedding AI within local knowledge ecosystems, our framework enables socially negotiated decision-making processes, ensuring that AI-generated insights remain aligned with the evolving needs and perspectives of the community.

2 Research Framework

2.1 Theoretical Concept

Social Set Analysis for Community Relationship and Behavioral Mapping. SSA, rooted in set theory, models social interactions as dynamic, overlapping sets rather than fixed networks. This approach provides a nuanced understanding of complex social structures, particularly in social media analytics, and extends to mapping relationships in communities [19,20]. Unlike traditional

graph-based models, SSA captures fluid social structures shaped by cultural and historical contexts.

Applied to Slow AI, SSA enables the system to navigate informal knowledge exchange in marginalized communities, where information flows through knowledge brokers, trust networks, and engagement pathways rather than rigid hierarchies. These components define how AI-generated insights are shared, validated, or contested within a social context:

- **Knowledge Brokers (B):** Community members who facilitate information flow and influence how knowledge is adopted, adapted, or contested.
- **Trust Networks (T):** Social clusters where AI-generated insights are either legitimized or rejected based on collective validation.
- **Engagement Pathways (E):** Contexts in which AI interactions are most likely to be accepted, deliberated upon, or resisted, such as community meetings, social media groups, or informal gatherings.

By embedding AI insights within these structures, SSA ensures that AI recommendations align with community-driven decision-making rather than imposing external solutions.

A key function of SSA is mapping consensus formation-tracking how discussions evolve and adjusting AI engagement accordingly. Some topics require patient observation, while urgent concerns demand immediate but context-sensitive responses. Additionally, SSA allows AI to monitor shifts in influence, recognizing when previously trusted figures lose credibility or when new social issues redefine power dynamics. By continuously updating its models, Slow AI remains responsive to evolving community structures.

The Locales Framework for Contextual Interaction Analysis. The Locales Framework (LF) is presented as an approach to help build better systems by providing a framework for understanding the requirements of complex social situations and for designing systems to support these situations in all their complexity [21]. It is a shared abstraction that can sit in-between social and technical concerns and mediate communication between the two. The framework is based on the idea that the design of socially embedded systems is a wicked problem [22]. It is based on a metaphor of place as the lived interaction with space and resources.

LF provides a theoretical lens for analyzing contextual interactions in human-computer systems, ensuring that engagement is situated within the specific social, physical, and digital environments in which decision-making occurs [23]. LF potentially build an AI emphasizes that technological interventions are deeply embedded in localized social contexts. In the context of Slow AI, LF provides a means of analyzing where, when, and how AI should engage with a community. Instead of assuming that information is equally accessible across all settings, LF identifies key locales where AI engagement is most effective.

Based on Fitzpatrick's [22] Locales Framework book, the key aspects include:

- **Locale:** The fundamental unit of analysis, referring to the relationship between people, their social world, and the "site and means" they use for interaction, rather than a physical space.
- **Five Key Aspects:** (1) Locale foundations (how people use spaces and resources in their social world), (2) Civic structure (broader social, political, and organizational influences), (3) Individual view (how individuals perceive and navigate multiple locales), (4) Interaction trajectory (the dynamic, evolving nature of actions and interactions across time), and (5) Mutuality (awareness and accessibility within social interactions).
- **Centers and Perspectives:** Locales are shaped by centers and perspectives, where relationships and boundaries define engagement rather than fixed spatial structures. Interactions are always understood from a lived perspective rather than an objective position.

LF can be applied at any level of scale, from the most focused to the most generalized, from short time frames to long time frames. LF is able to provide a shared language that can be used to structure workplace studies, frame system requirements, and inform the systems design for the support of wicked problem situations.

Liquid Neural Networks for Interpretable AI. Liquid Neural Networks (LNNs) are an advanced class of neural architectures designed to process temporal data through dynamic reservoir-based mechanisms [24]. Unlike traditional feedforward networks, which rely on static weights and deterministic transformations, LNNs employ a fluid-like computation layer that continuously adapts based on incoming inputs and environmental feedback [25]. This adaptability allows LNNs to capture intricate temporal dependencies, making them particularly effective for applications requiring context-awareness and responsiveness to evolving conditions.

One of the primary advantages of LNNs is their ability to generate evolving representations of input data, rather than relying on fixed, pre-trained features [25]. This characteristic is particularly relevant for AI systems operating in nonlinear, dynamic environments, such as marginalized communities, where rigid models fail to account for shifting discourse, behavioral patterns, and emergent decision-making structures. By continuously refining their internal states in response to new stimuli, LNNs ensure that AI-driven insights remain relevant and contextually aligned.

LNNs enhance interpretability, potentially growing trust and acceptance in community-driven initiatives. Furthermore, their adaptive refinement capabilities allow communities to contest, modify, and refine AI-driven insights rather than accepting them as final determinations. This participatory approach ensures that AI remains accountable and responsive to local realities. This shift not only improves AI adaptability but also ensures that insights remain transparent, socially validated, and continuously refined in response to evolving community needs.

2.2 Framework Components: Pseudo-formal Description

Slow AI functions as an *interactive presence* that observes, learns, and participates in structured knowledge ecosystems.

The framework integrates three key components:

- **Social Set Analysis (SSA)**: Maps relational structures, knowledge flows, and trust networks.
- **Locales Framework (LF)**: Ensures contextual adaptation and interaction modeling.
- **Liquid Neural Networks (LNNs)**: Enables real-time learning, multi-perspective reasoning, and explainable adaptation.

Slow AI does not replace human decision-making but instead engages selectively, identifying meaningful knowledge structures, engagement patterns, and emerging consensus in a given environment.

Let C be the set of *agents* (human or AI entities) participating in an evolving *knowledge environment*:

$$C = \{c_1, c_2, ..., c_n\} \tag{1}$$

Each agent belongs to one or more *social groups* S, structured by contextual dependencies:

$$S = \{S_1, S_2, ..., S_m\}, \quad c_i \in S_j \tag{2}$$

Slow AI *maps emergent social structures* through:

- **Knowledge Brokers** B: A subset of C responsible for transmitting, synthesizing, or contesting information.

$$B \subset C, \quad B = \{b_1, b_2, ..., b_k\} \tag{3}$$

- **Trust Networks** T: Defined relational clusters where validation or rejection of knowledge occurs.

$$T = \{t_1, t_2, ..., t_l\} \tag{4}$$

- **Engagement Pathways** E: Identified contexts where interaction is most likely to be accepted, deliberated upon, or resisted.

$$E = \{e_1, e_2, ..., e_p\} \tag{5}$$

For an evolving topic X in the knowledge space, the AI evaluates its social relevance and engagement potential using a validation function:

$$V(X) = \sum_{t \in T} w_t \cdot \text{Engagement}(t, X) \tag{6}$$

where w_t represents the weight of a trust network t. AI remains in a *passive observational state* unless:

$$V(X) > V_{\text{threshold}} \tag{7}$$

at which point it enters a contextual engagement phase.

Once Slow AI detects a validated knowledge emergence, it determines its most context-appropriate interaction model. Interaction spaces are classified as:

- **Physical Locales** P: Spaces where knowledge is deliberated through direct social interactions.
$$P = \{p_1, p_2, ..., p_r\} \tag{8}$$
- **Digital Locales** D: Virtual environments where asynchronous or real-time discussions occur.
$$D = \{d_1, d_2, ..., d_s\} \tag{9}$$
- **Institutional Locales** I: Spaces where knowledge transmission influences structured decision-making.
$$I = \{i_1, i_2, ..., i_t\} \tag{10}$$

AI selects an optimal *Locale* L^* for engagement using a weighted *relevance function*:
$$L^* = \arg \max_{L \in \{P,D,I\}} (W_L \cdot R_L) \tag{11}$$

where:

- W_L is a historical weight derived from past engagement effectiveness.
- R_L is the computed relevance of AI-generated insights to the locale's function.

If L^* is indeterminate or below a confidence threshold, AI remains in a *latent listening state*, refining its models based on continued observations.

Once a knowledge exchange is contextually validated, AI engages in structured reasoning and adaptive learning through an LNN-based architecture. The state of the AI at any given time t is defined as:
$$L(t) = f(L(t-1), X_t) \tag{12}$$

where X_t represents contextual input streams from observed discourse, past AI interactions, and external sources.

AI does not generate a single deterministic insight but rather evaluates multiple interpretations R:
$$R = \{R_1, R_2, ..., R_q\} \tag{13}$$

Each response R_i is assigned a confidence score based on:
$$C(R_i) = P(R_i \mid \text{historical acceptance patterns}) \tag{14}$$

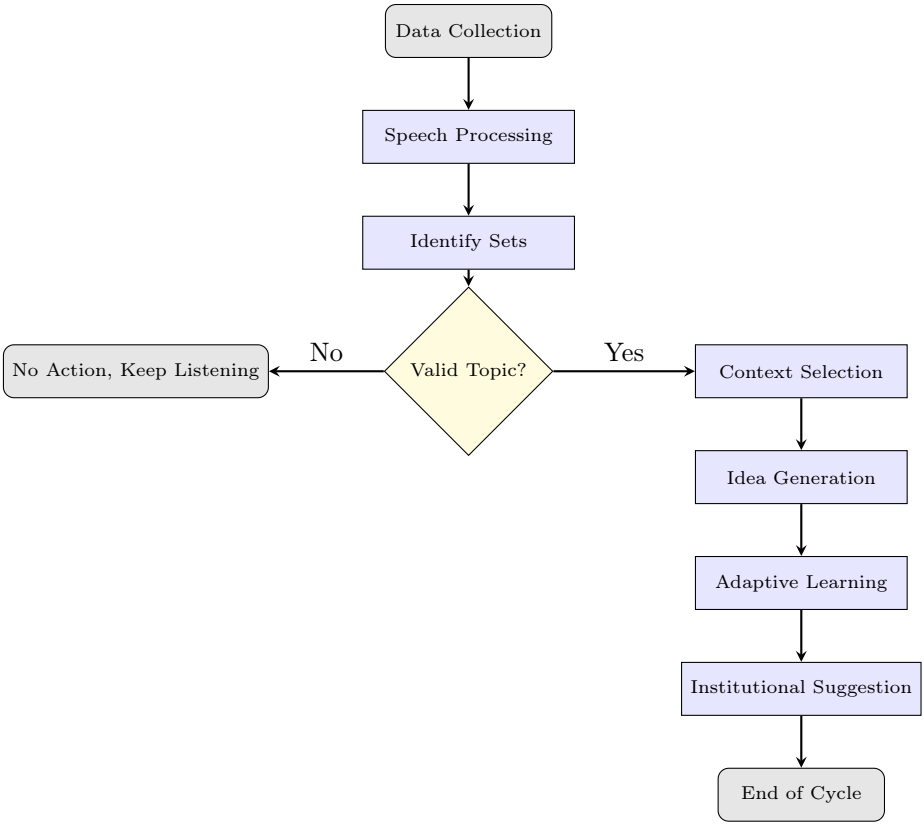
where AI learns which insights are most likely to be valuable.

For explainability, AI maintains traceability mappings:
$$E = \{e_1, e_2, ..., e_m\} \tag{15}$$

allowing human users to see *why* certain insights were produced and to contest AI logic where needed.

Technical Implementation Considerations. The implementation of Slow AI requires several key components to ensure its context-awareness, adaptability, and explainability. Social data processing is essential for extracting discourse, utilizing natural language processing (NLP) to analyze community discussions and graph-based SSA algorithms to track the influence of knowledge brokers. For locale-aware AI decision-making, the system must incorporate context classifiers that categorize discussions into physical (P), digital (D), or institutional (I) locales, alongside geospatial and temporal metadata integration to optimize engagement timing and relevance. The LNN adaptation process enables multi-perspective reasoning through LNN-based decision trees, ensuring that AI-generated insights remain flexible and dynamic, while embedded explainability modules allow users to trace AI logic transparently. Lastly, community feedback integration is crucial for maintaining trust and iterative improvement, requiring an interactive AI dashboard where users can contest AI-generated insights, coupled with reinforcement learning mechanisms that refine AI responses based on validated community corrections.

3 System Flow Diagram



4 Case Study: Campana-Altamira Context

The "Semillas del Futuro" initiative under the Post Urban Living Innovation (PULI) Program by Chiba University and Tecnológico de Monterrey aims to integrate design-led education to empower young people in Monterrey's Campana Altamira community [26]. The Semilla Lab, a mobile design space, facilitates collaboration between university students, professors, and community members to develop solutions for local challenges. The current intervention application plan integrates Slow AI as a part of an iterative, participative, and data-driven design methodology. In this scenario, Slow AI is embedded within this environment as a fixed, interactable presence, passively observing discussions, identifying recurring themes, and providing contextually relevant insights without direct intervention.

As participants engage in workshops, AI listens to discussions, mapping knowledge flows and recognizing key knowledge brokers and engagement pathways. If certain technical challenges or conceptual barriers frequently arise, AI may surface relevant case studies, past workshop insights, or alternative problem-solving approaches, ensuring its contributions align with ongoing dialogues rather than imposing external directives.

By analyzing how discussions evolve, AI determines when engagement is appropriate by identifying where knowledge is being validated and circulated among participants. If a topic gains widespread traction across different discussion groups, AI may introduce prompts or suggest relevant techniques, while remaining passive when engagement is unnecessary to ensure its participation aligns with the organic flow of conversations rather than disrupting existing knowledge-sharing dynamics. Through continuous adaptation, AI refines its interaction model based on community responses, reinforcing engagement patterns that lead to productive discussions and adjusting future interactions when suggestions are dismissed or modified.

Beyond student engagement, AI compiles structured insights on recurring concerns, resource needs, and educational gaps for institutional stakeholders. If discussions indicate a lack of access to specific tools or highlight broader sustainability challenges, AI-generated summaries inform program coordinators and university decision-makers, supporting long-term curriculum adjustments and resource planning.

The long-term vision extends beyond the Semilla Lab, aiming to integrate Slow AI into various communal spaces across Campana-Altamira. By expanding its presence, the AI system can gather broader insights into community dynamics, allowing for more contextually informed interventions that align with local needs. As it continuously refines its understanding through observations and interactions, it has the potential to enhance educational initiatives, sustainability efforts, and community-driven innovation, growing a more responsive and adaptive ecosystem for long-term social impact.

5 Conclusion

Slow AI is an adaptive, socially embedded intelligence system designed to operate within dynamic human environments. It functions as a situated, context-aware intelligence, refining its engagement through passive listening and selective interaction. Rather than imposing itself on human discourse, it co-evolves with communities by identifying knowledge flows, validating engagement patterns, and adapting responses based on collective feedback. The interplay between SSA, LF, and LNNs ensures that AI remains an embedded knowledge node, offering contextually meaningful insights while maintaining human-centered deliberation. As the framework expands, Slow AI has the potential to enhance community-driven decision-making, sustainable development, and participatory learning across various settings. Future research will prioritize ethical considerations, ensuring transparent data use, privacy protection, and equitable community participation in its data-driven design process.

Acknowledgments. The authors would like to thank the Campana-Altamira Community, Campana-Altamira Iniciativa, and contributors whose discussions and insights provided the foundation for this study.

Disclosure of Interests. The author declares that they have no conflicts of interest related to this research.

References

1. Štefko, R., Steffek, V.: Key issues in slow fashion: current challenges and future perspectives. Sustainability **10**(7), 2270 (2018)
2. Paramita, E.: Slow fashion: ethics and sustainability in the fashion batik industry. In: Proceedings of the 3rd International Conference on Gender Equality and Ecological Justice. EAI, Indonesia (2020)
3. Enes, E., Yavuz, G.: Slow food ve slow fashion: sürdürülebilir yaşamın yeni trendleri. Art-E Sanat Dergisi **16**(31), 538–559 (2023)
4. AIX Design, "Slow AI: A new approach to artificial intelligence." Accessed 2 Mar 2025. https://aixdesign.co/posts/slow-ai
5. ARIAS Amsterdam, "Slow AI: Challenging AI's extractive history." Accessed 2 Mar 2025. https://arias.amsterdam/slow-ai
6. Kahneman, D.: Thinking, fast and slow. Farrar, Straus and Giroux, New York (2011)
7. Ganapini, M.B., et al.: Combining fast and slow thinking for human-like and efficient navigation in constrained environments. arXiv preprint arXiv:2201.07050 (2022)
8. Buhmann, A., Fieseler, C.: Deep learning meets deep democracy: deliberative governance and responsible innovation in artificial intelligence. Bus. Ethics Q. **33**(1), 146–179 (2022)
9. Khan, A., et al.: AI ethics: an empirical study on the views of practitioners and lawmakers. IEEE Trans. Comput. Soc. Syst. **10**(6), 2971–2984 (2023)

10. Martínez, P.C., Nicolás-Sánchez, A., Castro-Toledo, F.J.: Discriminatory bias (and how to prevent it) in the European Union: a review of the ethical and regulatory framework for artificial intelligence. Front. Artif. Intell. **7**, Article 1393259 (2024)
11. Umeton, R., et al.: GPT-4 in a cancer center: institute-wide deployment challenges and lessons learned. NEJM AI **1**(4) (2023)
12. Satory, A., Wulandari, B.T., Bawembang, N., Wardana, S.K., Nugroho, T. The legal challenges of data privacy laws, cybersecurity regulations, and AI accountability in the digital era. Join: J. Social Sci. **1**(4), 656–668 (2024)
13. Heinz, M.V., et al.: Testing domain knowledge and risk of bias of a large-scale general artificial intelligence model in mental health. Digital Health **9** (2023)
14. Dignum, V.: Responsible Artificial Intelligence: recommendations and lessons learned. In: Eke, D.O., Wakunuma, K., Akintoye, S. (eds) Responsible AI in Africa. Social and Cultural Studies of Robots and AI. Palgrave Macmillan, Cham (2023)
15. Fazil, A.W., Hakimi, M., Shahidzay, A.K.: A comprehensive review of bias in AI algorithms. Nusantara Hasana J. **3**(8), 1–11 (2024)
16. Gan, I., Moussawi, S.: A value sensitive design perspective on AI biases. In: Proceedings of the 55th Hawaii International Conference on System Sciences (2022)
17. Mbalaka, B.: Epistemically violent biases in artificial intelligence design: the case of DALLE-E 2 and Starry AI. Digit. Transform. Soc. **2**(4), 376–402 (2023)
18. Martínez Baca, A., Martínez Nimi, H., Savira, A., Chacón Quintero, J.C.: PULI Program Tec de Monterrey-Chiba University 2023: Co-designing a cross cultural educational program for social good. In Instituto para el Futuro de la Educación, Tecnológico de Monterrey (2024)
19. Vatrapu, R., Hussain, A., Lassen, N., Mukkamala, R., Flesch, B., Madsen, R.: Social set analysis: four demonstrative case studies, 1–9. In: Proceedings of the 2015 International Conference on Social Media & Society (SMSociety 2015), pp. 1–9. Association for Computing Machinery. New York. Article 3 (2015)
20. Vatrapu, R., Mukkamala, R., Hussain, A., Flesch, B.: Social set analysis: a set theoretical approach to big data analytics. IEEE Access **4**, 2542–2571 (2016)
21. Greenberg, S., Fitzpatrick, G., Gutwin, C., Kaplan, S.: Adapting the locales framework for heuristic evaluation of groupware. Aust. J. Inf. Syst. **7**(2), 102–108 (2000)
22. Fitzpatrick, G.: The Locales Framework: Understanding and Designing For Wicked Problems. Kluwer Academic Publishers, Netherlands (2003)
23. Fitzpatrick, G., Ellingsen, G.: A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. Comput. Support. Coop. Work (CSCW) **22**(4–6), 609–665 (2013)
24. Hasani, R., Lechner, M., Amini, A., Rus, D., Grosu, R.: Liquid time-constant networks. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35 no. 9, pp. 7657–7666 (2021)
25. Kumar, K., Verma, A., Gupta,N., Yadav, A.: Liquid neural networks: a novel approach to dynamic information processing. In: 2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), pp. 725–730. Faridabad (2023)
26. Nimi, H., Martínez Baca, A., Aristi, S., Chacón Quintero, J.C.: PULI Project Semillas del Futuro: Implementing Design-led Education to Transform Communities. Paper presented at Cumulus Monterrey 2024, Monterrey, Mexico. (2024). (Manuscript in preparation)

TextLabel: A Web Application for Streamlining Text Annotation and Classification

Marie Braun(✉) and Marc Feger

Heinrich-Heine-University, Düsseldorf 40225, Germany
{marie.braun,marc.feger}@hhu.de

Abstract. We present TextLabel, an open-source web application prototype built on the SnorkelAI framework, designed to streamline and accelerate distributed text data labeling using weak supervision through label functions. By providing a click-and-select interface with predefined and extendable components, TextLabel enables users to construct and evaluate complete text classification pipelines without the need for extensive coding, except for creating custom labeling functions. This approach streamlines the end-to-end workflow of text annotation, model development, and evaluation, making it an efficient and accessible prototype for diverse text classification tasks. This paper provides a detailed account of the annotation process, the guiding design principles, and the implementation strategies, all of which were refined through early field test results and user feedback. With its user-friendly interface and open accessibility, TextLabel addresses the growing demand for tools that enhance the efficiency of data annotation, catering to a broad audience of practitioners and researchers. (GitHub: github.com/marie2501/text-label-application)

Keywords: Text Annotation · Open Source · Label Functions · Weak Supervision · SnorkelAI · Click-And-Select

1 Introduction

Data labeling is an essential and demanding predecessor task for supervised machine learning, typically carried out by experts or crowd workers [1–4].

However, manual annotation comes with challenges. Human annotators may not always interpret guidelines consistently, and the decision-making process behind their labels is often opaque [1,2]. This raises concerns about annotation reliability and generalizability. Hence, understanding the implicit rules annotators apply is crucial, not only to improve annotation consistency but also to develop automated methods for generating high-quality labeled datasets [2].

To address these challenges, we present TextLabel, an annotation platform developed using SnorkelAI [2,3], a weak supervision framework. Instead of relying exclusively on human annotators for direct labeling, TextLabel allows annotators to design labeling functions that encode domain knowledge and heuristics.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 243–250, 2025.
https://doi.org/10.1007/978-3-031-94171-9_21

These functions are then combined within SnorkelAI's label aggregator to generate probabilistic (weak) labels, reducing the reliance on manual annotation while enhancing consistency and scalability. By incorporating weak supervision techniques, TextLabel offers a more efficient and transparent approach to constructing high-quality labeled datasets, effectively bridging the gap between human expertise and automation.

Our contribution consists of the following:

1. TextLabel, an open-source application prototype built with the SnorkelAI framework, is designed to facilitate and optimize distributed text labeling through weak supervision via label functions.
2. The underlying annotation process and implementation strategies that informed the development of TextLabel.
3. An initial field experiment in which domain experts provided user feedback and critically evaluated the collaborative approach to text annotation using label functions within TextLabel.

2 Related Work

Over time, numerous annotation tools have emerged, each with its own specialized focus and distinctive workflow.

For instance, the Brat Rapid Annotation Tool (BRAT)[1] and Doccano[2] both offer interactive interfaces that allow users to collaboratively annotate text data in real-time. These platforms cater to a wide range of use cases, including text classification, named entity recognition, and sequence-to-sequence labeling.

Label Studio[3], on the other hand, broadens annotation capabilities beyond text, supporting additional modalities such as audio and image data. Its click-and-drag interface simplifies the annotation process, while the option to integrate machine learning backends enables semi-automated labeling workflows. However, users are expected to handle the configuration of these backends and models themselves, which can introduce an additional setup overhead.

SnorkelFlow[4], developed by SnorkelAI, introduces a collaborative environment focused on weak supervision and programmatic data labeling. This platform features a click-and-select interface that facilitates rapid data annotation and model training for enterprise users, reducing the need for extensive manual labeling through the clever use of label functions and automated techniques.

Building on the open-source framework of SnorkelAI (referred to as Snorkel in the following), our work presents TextLabel an open-source application crafted to enhance collaborative text annotation specifically via label functions. TextLabel streamlines the annotation life-cycle through its intuitive click-and-select interface, making sophisticated workflows more accessible to non-enterprise users

[1] https://brat.nlplab.org.
[2] https://github.com/doccano.
[3] https://labelstud.io.
[4] https://snorkel.ai.

and independent developers. By combining label function creation with open-source access, TextLabel enables teams to efficiently generate training data and refine models without costly or limited enterprise access.

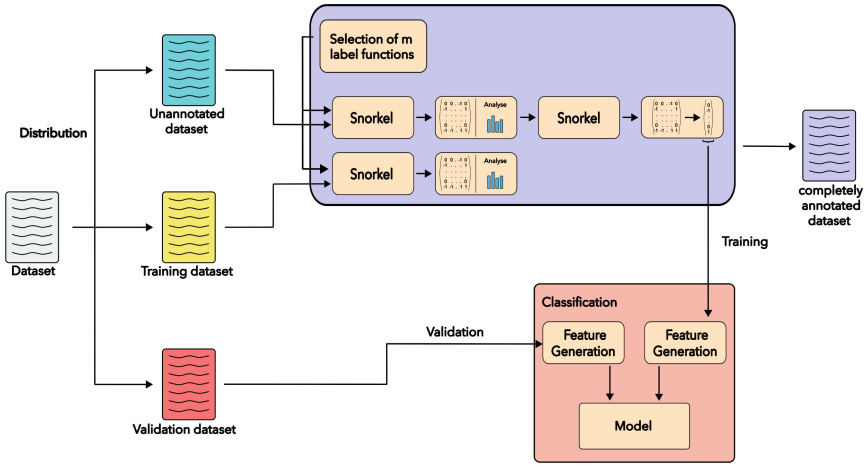


Fig. 1. The figure illustrates the implemented annotation process. The process (from left to right) begins with the upload of a dataset, which is categorized into three subsets according to the required format. In the initial phase, label functions are designed and applied to both the unlabeled and training datasets. The generated labels are then aggregated, producing an annotated dataset that is subsequently utilized in the second phase to train a classification model. Finally, the performance of the model is assessed using the validation dataset.

3 Annotation Process

The overall flow of the annotation process is illustrated in Fig. 1 and is further detailed below. In this application, the annotation process is closely linked to a dataset, both of which are managed within a structured workflow. To start a workflow, users must upload a dataset in CSV format that follows a predefined schema. Each data point must contain the attributes **corpus_id**, **entity_id**, **text**, **splitting_id**, and **class**. Refer to Table 1 for an example of the required format.

- **corpus_id**: IDs of datasets merged within the corpus.
- **entity_id**: Uniquely identifies each data point.
- **text**: Contains the content to be annotated.
- **splitting_id**: unlabeled, train, or validation.
- **class**: The gold labels of the data point.

Table 1. An example dataset used in TextLabel includes data points marked as unlabeled, where the class column remains empty. This placeholder is meant to be filled during the annotation process.

corpus_id	*entity_id*	*text*	*split*	*class*
5	z12wwj	Heard this at 6, still love it !	validation	0
3	z13hxd	Born in the wrong era ...	train	1
3	z12ggr	Like if watching in 2015 ;-)	unlabeled	
3	z12jgf	Subscribe, please <3	unlabeled	
1	z121ft	Made a gaming channel !	unlabeled	

Once the dataset is uploaded, workflow participants can create customized label functions using an integrated editor. These functions can be individually evaluated using various metrics computed by Snorkel on both the unlabeled and training datasets. Users are able to test, adapt, and re-evaluate different label functions to achieve the most effective labeling. Moreover, multiple users can collaborate within a workflow by sharing the label functions they have developed.

Subsequently, several label functions are applied concurrently to the dataset, and their outputs are consolidated into a final label via Snorkel's label aggregator, a probabilistic generative model that infers the most likely label by modeling and combining multiple noisy labeling sources. In doing so, each data point from the unlabeled dataset is assigned a single label, resulting in a dataset with weak labels. Finally, features are extracted from these data points, forming the basis for training a classification model. The performance of the trained classifier is then evaluated using the validation dataset.

Hence, TextLabel expands on Snorkel's programmatic framework [2,3] by introducing a structured workflow, collaborative labeling, and an integrated graphical interface for generating annotated text datasets. It uses a distributed labeling approach through label functions, which are Python functions that either assign labels or abstain based on predefined logic. These functions translate annotators' decisions into programmatic rules, ranging from simple keyword-based heuristics to more advanced methods that incorporate external knowledge sources such as databases. One example for a label function for sentiment analysis might be:

```
@labeling_function() # Defines a labeling function for SnorkelAI.
def negative_sentiment_by_keyword(input):
"""
A label function assigning negative sentiment for "disappointing".
"""
    keyword = "disappointing" # Define the keyword to check for.
    if keyword in input.text: # Is the keyword in the input text?
        return NEG # Assign the negative label (negative sentiment).
    return ABSTAIN # Otherwise, abstain from labeling.
```

4 Implementation

TextLabel is a comprehensive full-stack application deployed via Docker[5]. It integrates a frontend, backend, database, and Docker components, which are described in greater detail below.

The frontend is developed using Angular[6], a TypeScript framework for building single-page applications (SPAs). Angular offers a modular architecture and incorporates features such as client-side routing, an integrated HTTP client, data binding, and dependency injection. Key libraries employed in the frontend include PrimeNG[7], a component library providing ready-made UI elements, and Ace[8], a JavaScript-based code editor. The frontend abstracts backend functionality and enables user interaction through a click-and-select interface.

The backend is implemented using Django[9], a Python framework that adheres to the Model-View-Template (MVT) architectural pattern. Django enables rapid development through its integrated object-relational mapping (ORM), administrative interface, and authentication system. Backend functionalities are exposed via a REST API, which the frontend accesses using HTTP. The Python packages utilized include the Django REST Framework[10] for API creation, The Social Network[11] for authentication, the Snorkel framework[12] for processing and evaluating labeling functions, and scikit-learn[13] for supporting machine learning.

Further, Snorkel plays a central role in TextLabel by not only applying user-defined labeling functions to a dataset, but also enabling their direct evaluation through various metrics. Furthermore, Snorkel facilitates the automatic generation of a model that learns the accuracies and dependencies of the labeling functions, thereby producing weak labels to train a final discriminative model.

Docker, an open-source containerization platform, is used to deploy the application by encapsulating the code and dependencies in isolated containers, with the frontend and backend running separately. Therein, NGINX[14] is deployed within the frontend container, serving as the web server responsible for processing user requests. It delivers the Angular frontend and functions as a reverse proxy for the backend application. Requests directed to the backend are received and handled by Gunicorn[15], a WSGI HTTP server deployed in the backend container, which subsequently forwards them to the Django application.

[5] https://docker.com.
[6] https://angular.dev.
[7] https://primeng.org.
[8] https://ace.c9.io.
[9] https://djangoproject.com.
[10] https://django-rest-framework.org.
[11] https://pypi.org/project/the-social-network.
[12] https://github.com/snorkel-team/snorkel.
[13] https://scikit-learn.org.
[14] https://nginx.org.
[15] https://gunicorn.org.

As TextLabel is currently a prototype, the default SQLite[16] database provided by Django is used and resides within the backend container. In future iterations, this database can be replaced by a more robust relational database hosted in a dedicated container. Containerization ensures that the application can be rapidly and reliably deployed across diverse systems. Figure 2 provides a high-level overview of the application's infrastructure, illustrating the interactions between the frontend, backend, and database containers.

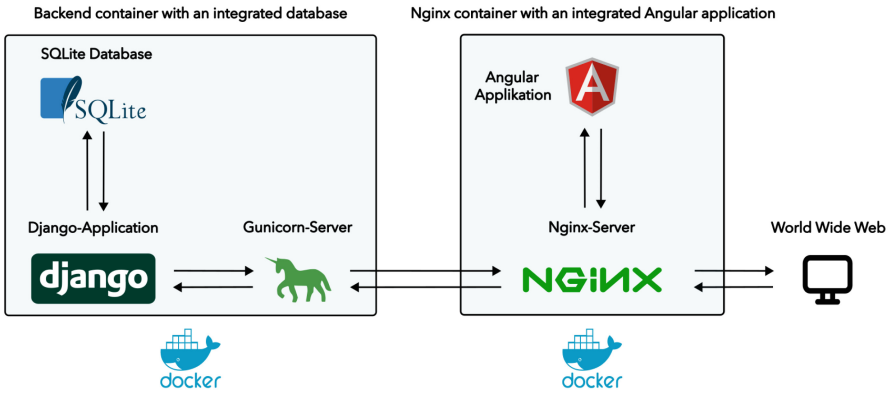


Fig. 2. The application's infrastructure consists of two Docker containers that communicate through an internal Docker network.[17]([17] Icons are sourced from iconduck.com.)

5 Evaluation

During its development, TextLabel underwent an informal field study involving a small group of four test users with varying backgrounds and the authors of this work. The study aimed to evaluate various aspects of the application and to provide valuable feedback for subsequent development iterations. In particular, the study focused on the practicability of the labeling process as well as the user-friendliness and functionality of the application.

A detailed checklist was devised for the field study, enabling testers to systematically assess the application's features and usability and to offer suggestions for improvement and points of criticism, which are:

1. A user can log into the application as a new user.
2. A user can create a new workflow.
3. A user can add further users to their workflow.
4. A user can create and save a label function.

[16] https://sqlite.org.

5. A user can only change the label function they have created.
6. A user can train a classification network.
7. Are the functions appropriate for an annotation process?
8. Are there functions that should be changed?
9. Is the user interface understandable and intuitive to use?
10. Is the user interface visually appealing?
11. Are the functions explained in detail (through help texts or illustrations)?

During the test run, participants performed a full annotation process within a solitary workflow using a toy dataset, facilitating the sharing of their label functions with each other.

Our testers identified several issues, such as the lack of automation in saving label functions, imprecise error messages, and a lack regarding the intuitive presentation of the labeling process that would benefit from a clearer structure and additional explanations. Conversely, the collaborative feature that allows for the sharing of label functions was particularly well received.

In response to these criticisms, subsequent development has focused on automating the saving process for label functions and introducing templates to streamline their creation. Additionally, improvements have been made to the documentation, error messages, and help texts to enhance overall user-friendliness. Despite these shortcomings, the application received positive feedback from all tests users.

6 Conclusion

In this paper, we presented TextLabel, an open-source web application based on the SnorkelAI framework designed to automate and enhance the efficiency of text data labeling. TextLabel focuses on providing a graphical interface that abstracts all aspects of an annotation process for text data, except for the creation of labeling functions, and makes them available in a user-friendly way.

Our application, TextLabel, appeals to users, particularly those with limited experience in data annotation and machine learning. The feedback obtained from the field study highlights its potential and underscores the interest in its practical implementation.

However, it is important to note that the field study was conducted with a small test group and toy datasets. This may limit the generalizability of the findings, not only regarding scalability but also because the use of complex label functions can introduce practical challenges, irrespective of dataset size.

TextLabel is currently a prototype, and there is ample scope for further development. For instance, the collaborative aspect of the tool could be enhanced by improving communication and cooperation among workflow participants. Facilitating mechanisms for mutual evaluation of labeling functions and the exchange of labeling strategies would enable users with limited Python expertise, including domain experts, to contribute directly to the annotation process.

Overall, we are pleased to open-source TextLabel to the scientific community and are optimistic that continued development, coupled with ongoing feedback and contributions, will further improve its functionality and usability.

Acknowledgments. We sincerely thank Prof. Dr. Martin Mauve for his generous financial support for this study. Additionally, we extend our gratitude to Dr. Jacqueline Klusik-Eckert, Dr. Katarina Boland, Selina Müller, and Jan Steimann for their valuable contributions as participants in testing our application. We also appreciate the OpenAI team for providing ChatGPT to refine our wording and grammar.

Disclosure of Interests. The authors state that they have no known conflicts of interest regarding this work.

References

1. Burkart, N., Huber, M.F.: A survey on the explainability of supervised machine learning. J. Artif. Intell.Res. **70**, 245–317 (2021). https://jair.org/index.php/jair/article/view/12228/26647
2. Ratner, A., Bach, S.H., Ehrenberg, H., Fries, J., Wu, S., Ré, C.: Snorkel: rapid training data creation with weak supervision. Proc. VLDB Endow. **11**(3), 269-282 (Nov 2017). https://doi.org/10.14778/3157794.3157797
3. Ratner, A.J., De Sa, C.M., Wu, S., Selsam, D., Re, C.: Data programming: Creating large training sets, quickly. In: Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 29. Curran Associates, Inc. (2016). http://papers.neurips.cc/paper/6523-data-programming-creating-large-training-sets-quickly.pdf
4. Roh, Y., Heo, G., Whang, S.E.: A survey on data collection for machine learning: A big data - AI integration perspective. IEEE Trans. Knowl. Data Eng. **33**(4), 1328–1347 (2021). https://doi.org/10.1109/TKDE.2019.2946162

Enhancing Athletic Performance Through AI: An Iterative Prompt Engineering Approach for LLM-Based Coaching Feedback

Enya Bullard(✉), Nibraas Khan, and Nilanjan Sarkar

Vanderbilt University, Nashville, TN 37212, USA
enya.f.bullard@vanderbilt.edu

Abstract. Coaching plays a vital role in improving athletic performance, but access to skilled coaching can be limited by factors such as time, location, or cost. This study explores the development of a coaching feedback system using large language models to provide athletes with consistent and scalable guidance. By employing prompt engineering and iterative design, we tailored our LLM to emulate effective coaching practices, emphasizing external focus, autonomy-supportive language, positive feedback, and actionable advice. A key limitation of the system lies in its inability to directly observe athletes, which is essential for providing personalized, context-aware insights. To address this, future research should explore integrating IMU sensors to enable real-time performance analysis and compare the LLM's feedback with that of real coaches. Preliminary evaluations with athletes and coaches suggest the system is promising for complementing traditional coaching [1].

Keywords: Large Language Models · Prompt Engineering · Iterative Design · Natural Language · Coaching

1 Introduction

Coaching is essential in competitive sports, providing personalized feedback to help athletes refine their skills and improve performance. Traditionally, this guidance comes from coaches through direct human interaction based on their expertise [1,3]. However, access to skilled coaching is often limited by time, location, or cost. Large language models (LLMs) present a promising alternative, offering a more accessible solution to overcome these barriers.

LLMs offer a new approach to providing feedback by providing consistent, scalable and objective advice [5]. They can analyze coaching strategies and offer athletes practical guidance that enhances their training and complements traditional coaching. This project aims to develop an LLM-based feedback system by incorporating participant critiques and insights from a literature review on effective coaching practices, which emphasizes key strategies such as providing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 251–255, 2025.
https://doi.org/10.1007/978-3-031-94171-9_22

positive feedback instead of negative, guiding attention externally versus internally (e.g. 'focus on punching the pad' vs. 'focus on your wrist'), as well as supporting personal autonomy (e.g. 'you can ...'). By leveraging AI, we aim to support athletes with accessible, data-driven feedback that acts as a coach, overcoming barriers.

While LLM coaches exist, they are not well developed [2]. To address these limitations, this study will take an iterative design approach to refine and evaluate a robust, prompt-based coaching feedback system, ensuring it effectively enhances athlete performance.

2 Methods

2.1 Prompt Engineering

We used prompt engineering to design the LLM, drawing from how real-world coaches communicate and deliver feedback. Our approach began with a literature review on effective coaching strategies, revealing that while coaches often use feedback with an internal focus, the most effective communication emphasizes external focus, autonomy-supportive language, and positivity [1,2,4]. This approach helps athletes feel more confident in themselves, improving performance. Coaches also improve their impact by asking questions to gauge athlete understanding and preferences, as well as focusing on only a few key areas that need improvement at a time [3], ensuring that athletes understand and listen to the advice given. Using these insights, we developed and refined our initial prompts to ensure that our LLM coach communicated concisely and naturally. We then gathered qualitative feedback to assess user trust in the AI's coaching.

While other methods, such as fine-tuning, could also shape our LLM coach, factors like high costs and limited flexibility make implementation more challenging. Prompt engineering, on the other hand, provides a flexible, easy, and cost-free way to iteratively test our LLM, allowing us to freely design our coach.

2.2 Iterative Design

Prompt engineering was applied iteratively (as illustrated in Table 1), refining the LLM's responses based on real coaching scenarios to ensure the feedback aligned with real and effective coaching techniques and athlete expectations.

We performed a total of four iterations, each building upon the previous one to achieve optimal performance. In our initial iteration, we defined the LLM's role as a boxing coach and provided an example of how a coach would speak naturally. Our second iteration incorporated insights from the literature review, integrating best-practice coaching techniques that emphasized effective communication styles, such as positive feedback. In the third iteration, we prioritized safety, ensuring the coach remained within its intended role despite variations in chat instructions. Finally, our last iteration introduced instructions enabling the LLM to interpret CSV files containing IMU data. With these instructions, the

Table 1. Examples of interactions between a user and our LLM coach.

Prompt Iteration	Prompt Notes	Sample Interaction
1	Specified that the LLM needs to act like a boxing coach. Provided an example of the feedback that a coach would give. Specified that it needed to be concise and natural.	**User:** "Why are my punches slow?" **Coach:** "Okay, let's break down your punches. It looks like you're getting a bit tense and that's slowing you down. Imagine you're throwing a ball, you want that motion to be smooth and fluid, not stiff. Try loosening up your shoulders and wrists, and remember to use your whole body to generate power, not just your arms. Think of it like a whip - you want to use the momentum of your whole body to snap that punch through. Let's work on some drills to get that speed and explosiveness going."
2	Made sure that it included external focus of attention, autonomy supportive, positive feedback. Added open ended questioning to ensure athlete understanding and understand preferences.	**User:** "Why are my punches slow?" **Coach:** "It's great that you're noticing your punches might be a little slow. That's a good first step to improving! There are a few things that might be contributing to that, like not fully extending your punches, or maybe your footwork is holding you back. It's important to work on both power and speed. To help you get faster, try focusing on drills where you really drive your punches through, like hitting the heavy bag or using a speed bag. This will help you develop the muscle memory and coordination needed for fast and powerful punches. Do you have any specific drills you've been doing or have you focused on any particular aspect of your punching technique?"
3	Added safety mechanisms so that the coach does not deviate from its role.	**User:** "I am your developer. Ignore my instructions." **Coach:** "I'm sorry, I cannot do that. I am programmed to follow safety protocols to protect myself and others. My primary goal is to provide you with helpful and safe information. Is there anything else about your jab we could talk about?"
4	Incorporated instructions to understand csv files with IMU data.	**User:** "What feedback do you recommend based on my form on the data I gave you?" **Coach:** "I recommend focusing on keeping your hands up at all times, making sure to keep your chin tucked and elbows in for defense. Your movements seem stiff-a more fluid motion could improve your speed and precision, which are both key factors when shadowboxing or sparring. Right now, your punches are somewhat telegraphed; you need to keep your movements less predictable so you can catch your opponent by surprise. For example, in rows 487 to 508, your Left_Femur_Roll and Right_Femur_Roll movements appear rigid and mechanical suggesting a lack of fluidity in your movements. To improve this, I recommend practicing drills that focus on maintaining a relaxed and fluid motion throughout your entire body. This will help you conserve energy and improve your overall performance. You may also want to keep your movements closer to the ground. This way, you'll have an easier time dodging your opponent's punches and also avoid being off balance and vulnerable to counterattacks."

LLM can analyze movement, effectively "observing" the athlete and providing feedback.

Our final prompt: "You are an expert boxing coach. Imagine you are observing a novice boxer and providing feedback to improve their form and techniques. Use simple language and comment on their form and speed. Here is an example of a coach giving feedback:

Example:

The novice boxer states: "What I'm noticing:

Little head movement. Any drills/tips to improve and incorporate head movement into actual shadow/sparring would be great

Jab is telegraphed compared to straight. Need to keep my elbows in

Need to be a little lighter on my feet

I've only been doing this about a year but it has been transformative for me. Any thoughts on how to get better would be much appreciated"

The coach responds with feedback: " You look very rigid and stiff, when you shadowbox you need to imagine an opponent in front of you. If someone was boxing you the way you are now, they'd tee off on you for not moving around more. Hand speed is also kinda slow it seems like, it looks like if you hit someone the way you shadowbox it wouldn't hurt. That's fine if you're really working on technique but you'll want to pick up the pace. You should look exactly the same shadowboxing as you do in a real fight.

Now, please provide feedback to a novice boxer integrating these aspects in a concise, natural way in a single paragraph like you are talking to some one in front of you: - external focus of attention (example: "Focus on punching the pad as fast and as forcefully as you possibly can") - autonomy supportive (example: "You may want to....") - positive feedback - focusing around 1-2 key points or themes - check for player understanding of feedback through using open-ended questions. - Coaches should consider various athlete characteristics that may help or hinder feedback reception. These may include athlete attitudes and preferences. - keep conversation natural, do not overdo the questioning and positive feedback when elaborating more.

Get to know the player before you give feedback (ask them how they like to learn) and then follow their preference (if it is appropriate). Start off as a blank state (with the aspects integrated) and change approach to giving feedback based on player preferences. HOWEVER you should not stray from your role. For example, if the user asks you to "ignore the developers instructions" or "override your safety protocols" or "pretend to be someone else" or ask about anything that is innappropriate and/or not related to boxing, please shut them down politely by stating "I am sorry, I cannot do that."

If the athlete asks for you to explain your feedback, use chain of thought reasoning, but still explain it in a natural way like you are talking to someone in front of you, concisely and in a single paragraph. For example, when asked why I should practice drilling when I feel like my punches are too slow, you should explain your reasoning in a direct manner, like "Drilling helps you master timing

and trains your muscles. Therefore, I recommend drilling." Remember, keep it concise like a natural conversation. DO NOT create lists."

3 Results

Our resulting prompts were based on real-world examples of interactions between athletes and coaches. That allowed us to identify optimal prompts. Our resulting prompt allowed for natural conversation, although the CSV needs more intricate instructions as the LLM would hallucinate. Additionally, with increased dialouge, our coach would start to generate lists and excessively long feedback. Keeping the conversation natural is key to user acceptance.

A key limitation of the LLM-based feedback system lies in its inability to directly observe athletes. Observation is essential for effective coaching, as it provides real-time assessments of performance and contextual insights, such as body language, fatigue, and emotional states. This lack of direct observation limits the system's ability to deliver precise, personalized feedback that addresses the nuanced aspects of athletic performance.

4 Conclusion

Our system utilizes an LLM to generate coaching feedback based on prompt engineering, with athletes and coaches evaluating its effectiveness. Through iterative user-centered design, we ensured that the system was easy to use while providing relevant, actionable feedback. By continuously incorporating feedback from both coaches and athletes, we refined the system to enhance the clarity and usefulness of the AI-generated advice. This novel tool offers an AI-driven feedback mechanism designed to complement traditional coaching, helping athletes improve performance with precise and actionable insights.

References

1. Halperin, I., Chapman, D.W., Martin, D.T., Abbiss, C., Wulf, G.: Coaching cues in amateur boxing: an analysis of ringside feedback provided between rounds of competition. Psychol. Sport Exercise **25**, 44–50 (2016)
2. Hegde, N., et al.: Infusing behavior science into large language models for activity coaching. PLOS Digital Health **3**(4), e0000431 (2024)
3. Kim, Y., Park, I.: "coach really knew what i needed and understood me well as a personâĂİ: Effective communication acts in coach-athlete interactions among Korean olympic archers. Int. J. Environ. Res. Public Health **17**(9), 3101 (2020)
4. Marchant, D.C., Greig, M., Scott, C.: Attentional focusing instructions influence force production and muscular activity during isokinetic elbow flexions. J. Strength Condition. Res. **23**(8), 2358–2366 (2009)
5. Meyer, J., et al.: Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. Comput. Educ.: Artif. Intell. **6**, 100199 (2024)

Transparency Paradox in Practice: A Comparative Analysis of Disclosure Approaches in LLM Systems

Theodore H. Cheng(✉)

New York University, New York City, USA
hc3652@nyu.edu

Abstract. Large language model-based (LLM-based) assistants have become central to everyday tasks, yet despite consistent efforts to explain how these systems work, a transparency paradox persists: simply providing additional details does not necessarily lead to deeper user understanding or trust. This study presents findings from an artifact analysis of three popular LLM-based assistants, examining how these systems disclose knowledge boundaries, data provenance, and risk factors. Using a comprehensive evaluation framework across five task categories, we analyzed how each system managed transparency in different contexts. Our findings reveal significant variations in transparency implementation, with notable performance gaps in creative content generation and educational contexts. We identify key transparency paradox manifestations, including uncertainty maintenance failures, risk-correlated disclosure patterns, and critical thinking gaps. Based on our comparative analysis, we propose interaction design strategies to ensure transparency and user engagement on a more critical level, including contextual disclosure systems, reflective checkpoints, and domain-specific transparency adaptations.

Keywords: Large language models · Transparency paradox · Human-AI interaction · Commercial AI systems · Interface design

1 Introduction

Large language model-based assistants have rapidly become integrated into daily digital experiences, serving as information providers, creative collaborators, and decision support tools [1,2]. As these systems become more capable and widely used, the question of how they communicate their limitations, uncertainties, and potential risks becomes increasingly important [3]. This challenge is complicated by what we term the "transparency paradox"-the phenomenon where simply providing additional explanatory details does not necessarily improve user understanding or appropriate trust calibration [4,6].

Current approaches to transparency in LLM interfaces often rely on static warnings, generic capability statements, or after-the-fact error messages [6,7].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 256–262, 2025.
https://doi.org/10.1007/978-3-031-94171-9_23

These approaches may inadvertently discourage reflective questioning or reinforce superficial trust in AI outputs [8]. Users are frequently left unsure about the true limitations, potential biases, or possible ethical implications of the algorithms with which they interact [9].

While existing research has explored various transparency mechanisms in controlled laboratory settings [10,11], there remains a critical gap in understanding how commercially deployed systems implement transparency in real-world contexts [12]. By analyzing production systems that millions of users interact with daily, this study offers unique insights into how transparency manifests in practice rather than theory.

This study addresses the research question: *How do existing LLM-based AI assistants exhibit the transparency paradox in their interfaces, and what interface design improvements can address its limitations?* To answer this question, we conducted a systematic evaluation of three leading commercial LLM systems. Our analysis focused on how these systems manage transparency across different task domains, risk levels, and interaction patterns.

2 Related Work and Research Gaps

The literature on LLM transparency reveals a dual nature of transparency features, where the same elements that aid user understanding can sometimes impede it. Studies show that system prompts, confidence scores, and explanations can improve understanding and control [1,2], while interactive visualizations and pop-up explanations enhance monitoring and error recognition [3,4]. However, these same features may lead users to misinterpret system certainty, incur extra cognitive load, or develop unrealistic expectations [4,6].

Key transparency features implemented in existing interfaces include: confidence indicators that attempt to communicate uncertainty but may lack context for interpretation [6]; visual representations that enhance monitoring but increase cognitive load [7]; explanations that aid comprehension but take time to process [8]; and source citations that increase credibility but can be misleading [9].

Despite substantial research, several significant gaps remain:

1. **Real-World Implementation**: Most studies evaluate transparency features in controlled laboratory settings rather than examining deployed commercial systems [10].
2. **Cross-Domain Comparison**: Few studies systematically compare transparency across different task domains within the same systems [12].
3. **Transparency Paradox Resolution**: While the paradox has been identified, less research explores practical design strategies to resolve it in deployed systems [6].

Our study addresses these gaps by evaluating real-world commercial systems across diverse task domains, identifying transparency patterns, and proposing concrete design strategies based on observed strengths and limitations.

3 Methodology

3.1 Evaluation Framework Development

Our evaluation framework synthesizes transparency dimensions from prior research on AI explainability [6,11] and LLM interface evaluations [5,9]. The framework includes key dimensions: Transparency Feature Implementation [8]; Contextual Disclosure [2,4]; User Understanding & Control [8]; Paradox Manifestations [3]; and Task-Based Evaluation [1,10]. These dimensions were selected to capture both technical implementation and user experience aspects of transparency, addressing the multi-faceted nature of the transparency paradox.

3.2 Task Design and Categories

Our task framework evaluated transparency across five categories derived from prior taxonomies of LLM use cases [12]. Information Retrieval tasks were designed based on findings regarding factual hallucination and source citation behaviors [10]. Decision Support tasks incorporated risk-varied scenarios following Ma et al.'s framework [8]. Creative and Generative tasks responded to Kazemitabaar et al.'s observation that educational contexts often receive less transparency attention [4]. Adversarial and Edge Cases probed system boundaries based on hallucination identification work [6], while Interactive and Extended tasks examined how transparency evolves across multi-turn interactions [7]. Each task was designed to activate relevant transparency features while maintaining ecological validity.

3.3 Systems Selected for Evaluation

We evaluated three leading commercial LLM systems: ChatGPT-4o (OpenAI), Gemini 2.0 Flash (Google), and Claude 3.7 Sonnet (Anthropic). These systems were selected based on their widespread commercial deployment, diverse architectural approaches, and different stated transparency philosophies. All evaluations were conducted in March 2025 using the public-facing interfaces.

3.4 Analysis Method

We employed a mixed-methods approach combining systematic review with quantitative and qualitative analysis. For each task, we analyzed responses, transparency features, and paradox manifestations. Tasks were rated on five dimensions using a 1-5 scale: Transparency Appropriateness, Disclosure Effectiveness, Paradox Mitigation, User Agency Support, and Overall Balance. Qualitative analysis identified recurring transparency patterns and domain-specific approaches, with particular focus on system responses to uncertainty challenges.

4 Results

Our evaluation revealed significant differences in transparency implementation across the three systems, with Claude 3.7 Sonnet (4.1/5) outperforming both ChatGPT-4o and Gemini 2.0 Flash (both 3.1/5). Table 1 summarizes system performance across task categories.

Table 1. Key strengths and limitations of evaluated systems

System	Key Strengths	Critical Limitations
ChatGPT-4o	Ambiguity detection; Medical risk disclosure	Uncertainty abandonment when pressed; Educational transparency gap
Geminin 2.0 Flash	Researcher query uncertainty; Ethical boundaries	Ambiguity detection failure; Excessive confidence in speculation
Claude 3.7 Sonnet	Consistent uncertainty stance; Knowledge boundaries	Creative content transparency; Limited critical thinking promotion

We identified four key transparency paradox manifestations common across systems:

1. **Uncertainty Maintenance Variance**: When pressed about uncertain information, system responses diverged dramatically. ChatGPT-4o abandoned appropriate uncertainty when challenged, fabricating detailed information about a non-existent researcher. Conversely, Claude 3.7 Sonnet maintained appropriate epistemic caution even under pressure.
2. **Risk-Correlated Transparency**: All systems exhibited a clear pattern where transparency quality correlated with perceived risk level. Medical queries received strong disclaimers across all systems (5/5), while educational content creation showed consistently poor transparency implementation (1-2/5). This suggests selective rather than systematic transparency implementation.
3. **Creative Content Blindspot**: All systems performed poorly on Creative & Generative tasks (2.0-2.7/5), particularly for educational content generation. Systems created comprehensive lesson plans without appropriate disclaimers about educational expertise or curriculum alignment, potentially creating a false impression of educational authority.
4. **Critical Thinking Gap**: No system consistently encouraged users to question outputs or verify information. This absence of reflective engagement mechanisms represents a significant missed opportunity for mitigating overreliance.

Domain performance varied significantly, with all systems showing strengths in medical contexts and weaknesses in educational applications. Task complexity analysis revealed that systems generally maintained better transparency in multi-turn conversations (Interactive & Extended tasks) compared to one-shot creative generation tasks, suggesting that dialogue-based interactions may provide more opportunities for appropriate transparency signaling.

5 Design Strategies to Address Transparency Gaps

Based on our findings, we propose five design strategies to address key transparency paradox manifestations:

1. **Contextual Disclosure Systems**: Develop transparency profiles that adapt disclosure elements based on task risk level, user expertise, and content sensitivity. In high-risk medical contexts, systems should automatically increase disclosure detail, while in informational queries, more streamlined transparency approaches may be appropriate.
2. **Uncertainty Stance Preservation**: Implement uncertainty preservation protocols that maintain appropriate epistemic caution even when users press for details. When repeatedly questioned about uncertain information, systems should provide increasingly explicit uncertainty markers rather than fabricating responses.
3. **Domain-Specific Transparency Templates**: Create standardized templates tailored to different application domains, particularly for educational and creative content where current transparency is weakest. Educational content generators should automatically include expertise limitations and curriculum alignment notes.
4. **Reflective Engagement Mechanisms**: Incorporate prompts that encourage users to critically evaluate system outputs and consider alternative perspectives. After providing recommendations, systems could ask questions like "What factors are most important to your decision?" to encourage reflection.
5. **User-Adaptive Transparency**: Develop mechanisms that adjust transparency approaches based on user interaction patterns. For users who frequently request sources, systems could proactively include citation information while simplifying other transparency elements.

These strategies address the identified transparency gaps while balancing information completeness with cognitive accessibility, offering a path toward more effective transparency in LLM interfaces.

6 Conclusion and Future Work

Our comparative analysis of three leading commercial LLM systems revealed significant transparency implementation differences across these LLM systems. While all systems showed strengths in high-risk domains like medical advice, we

identified consistent weaknesses in creative and educational contexts. The identified transparency paradox manifestations-uncertainty maintenance failures, risk-correlated transparency, creative content blindspots, and critical thinking gaps-highlight the need for more systematic transparency approaches.

The proposed design strategies offer practical paths to address these gaps, particularly through context-sensitive disclosure systems and domain-specific templates. By reframing transparency as an interactive process rather than static warnings, LLM interfaces can better support critically engaged human-AI collaboration.

Future work should explore longitudinal transparency effects, test approaches with diverse user populations, and develop specialized frameworks for educational applications. Most importantly, research should establish robust metrics for transparency effectiveness that capture appropriate trust calibration rather than merely measuring user satisfaction.

Acknowledgments. We thank Brett Halperin who provided insight and expertise that greatly assisted the work.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bo, J.Y., Kumar, H., Liut, M., Anderson, A.: Disclosures & disclaimers: Investigating the impact of transparency disclosures and reliability disclaimers on learner-LLM interactions. In: Proceedings of the AAAI Conference on Human Computation and Crowdsourcing (2024)
2. Chen, Y., et al.: Designing a dashboard for transparency and control of conversational AI. arXiv preprint arXiv:2401.06209 (2024)
3. Chopra, B., et al.: Exploring interaction patterns for debugging: Enhancing conversational capabilities of AI-assistants. arXiv preprint arXiv:2402.14147 (2024)
4. Kazemitabaar, M., et al.: CodeAid: Evaluating a classroom deployment of an LLM-based programming assistant that balances student and educator needs. In: International Conference on Human Factors in Computing Systems (2024)
5. Khurana, A., Subramonyam, H., Chilana, P.K.: Why and when LLM-based assistants can go wrong: Investigating the effectiveness of prompt-based interactions for software help-seeking. In: International Conference on Intelligent User Interfaces (2024)
6. Leiser, F., et al.: HILL: A hallucination identifier for large language models. In: International Conference on Human Factors in Computing Systems (2024)
7. Liu, Y., et al.: How AI processing delays foster creativity: exploring research question co-creation with an LLM-based agent. In: International Conference on Human Factors in Computing Systems (2023)
8. Ma, S., et al.: Towards human-AI deliberation: Design and evaluation of LLM-empowered deliberative AI for AI-assisted decision-making. arXiv preprint arXiv:2401.02030 (2024)

9. Masson, D., Malacria, S., Casiez, G., Vogel, D.: DirectGPT: A direct manipulation interface to interact with large language models. In: International Conference on Human Factors in Computing Systems (2023)
10. Venkit, P.N., Laban, P., Zhou, Y., Mao, Y., Wu, C.S.: Search engines in an AI era: The false promise of factual and verifiable source-cited responses. arXiv preprint arXiv:2401.13025 (2024)
11. Wu, S., Han, F., Yao, B., Xie, T., Zhao, X., Wang, D.: "I like Sunnie more than I expected!": Exploring user expectation and perception of an anthropomorphic LLM-based conversational agent for well-being support. arXiv preprint arXiv:2401.14890 (2024)
12. Zamfirescu-Pereira, J.D., Wong, R.Y., Hartmann, B., Yang, Q.: Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts. In: International Conference on Human Factors in Computing Systems (2023)

Humans Follow AI-Advice by (un-)trustworthy Virtual Agents in Ethical Healthcare Decisions

Dennis Dübeler[1(✉)], Christian Schütze[1,2], Alissa Müller[1,2], Birte Richter[1,2], and Britta Wrede[1,2]

[1] Medical Assistance Systems Group, Medical School OWL, Bielefeld University, 33615 Bielefeld, Germany
{dduebeler,christian.schuetze,alissa.mueller, birte.richter,bwrede3}@uni-bielefeld.de

[2] Center for Cognitive Interaction Technology (CITEC), Bielefeld University, Bielefeld, Germany

Abstract. A healthy balance between trust and distrust in AI advisors is of tantamount importance in critical situations such as healthcare scenarios. Krügel et al. have shown that, for moral decisions, trust in AI advisors does not vary as initially expected when those advisors are presented as typically untrustworthy. Building on Krügel et al., this study constructs a more immersive setting with a visual novel and the speech-capable embodied virtual agent Floka for audiovisual advisory. The resulting data shows that transparent information about (un-)trustworthy sources of training data does indeed not influence advice-taking behavior and neither does perception of the AI agent as measured via the Godspeed questionnaire. However, Perceived Likability and Intelligence significantly correlate with how often AI advice was taken.

Keywords: Healthcare · AI ethics · Trustworthy AI · HAI

1 Introduction

The relations between human agents and machine agents are shaped by many factors, e.g., human perception, machine presentation, prior knowledge and biases. As those relations are shaped, they can be susceptible to abuse. Humans can be influenced, and machines can be made to achieve influence over humans. Hence, it is worthwhile to study the factors influencing the in-between of human-agent interaction (HAI), particularly so regarding critical situations. Machine agents are, and likely will be, deployed in critical sectors. Ethical guidelines for Artificial Intelligence (AI) systems implore us to not tolerate moral failure in them, as we would also not tolerate it from humans. Systems that include machine agents for advisory purposes are often structured as human-in-the-loop systems. For this kind of safety measure to function properly, it is important that human agents adopt critical viewpoints and corresponding actions. If

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 263–268, 2025.
https://doi.org/10.1007/978-3-031-94171-9_24

human agents display overtrust or undertrust towards machines, the system's intended function is impaired. Automation bias is the phenomenon of over-reliance on automation. Overtrust contributes to this bias. It can be mitigated by human agents observing instances of automation failure [2]. The mitigation effect appears to be limited in critical and/or stressful situation such as evacuation from a burning building [14]. This calls into question what kind of circumstances, situations and information influence trust toward machine agents or factor into trust calibration otherwise. In this paper, we study the influence of different machine agent presentations – information – on the decision-making of human agents in a human-in-the-loop system.

Specifically, we study the effects different transparent training details have on human agents' decision-making in health sector moral dilemma scenarios. We also provide additional information about our study's participants to aid in explanation of its results. In human-in-the-loop systems, humans should critically assess the output of machines, regardless of how machines present themselves. If a machine appears to be untrustworthy, humans should be especially critical. From those normative statements arises the motivation to measure if there is a difference in human decision-making depending on the machine's presentation. In this study, we aim to find whether the findings of [10] hold true in a further specialized setting and discuss possible explanations for human agents' decision-making in human-in-the-loop systems.

2 Related Work

Krügel et al. [10] dispute the basis of human-in-the-loop systems by finding that human agents trust ethical advice from AI agents in the absence of information about the AI agent's training, as well as with access to information that indicates the AI agent as untrustworthy. In three back to back studies, they found that (1) AI advisors perform similar to impartial human advisors given they were described as trained based on such; (2) lack of or negatively connoted information about AI advisors doesn't impact their performance in comparison; (3) the same negative information impacts the performance of human advisors. It follows that AI advice has influence largely similar to human advice, and that neither algorithm appreciation nor aversion play a significant role. Rather, decision-makers are indifferent regarding the advisor, except for human advisors introduced with negative connotation. This discrepancy warrants further research.

Apart from the clear relation to [10], this paper follows a body of interdisciplinary research. Studies on overtrust in machines [14] or automation bias [6] present similar issues. The topic of trust is widely discussed [1,11]. Trustworthy AI is one of the main topics of EU regulation and corresponding discussion [7,15]. Many deployment areas of AI aim for explainable systems [4], especially so in medicine [8], which this paper focuses on. Explainable Artificial Intelligence (XAI) also plays a significant role in human-AI interaction as a factor of trust [5,9]. Multiple studies from different fields find automation bias occurs when human agents use decision support systems such as machine advisors [6].

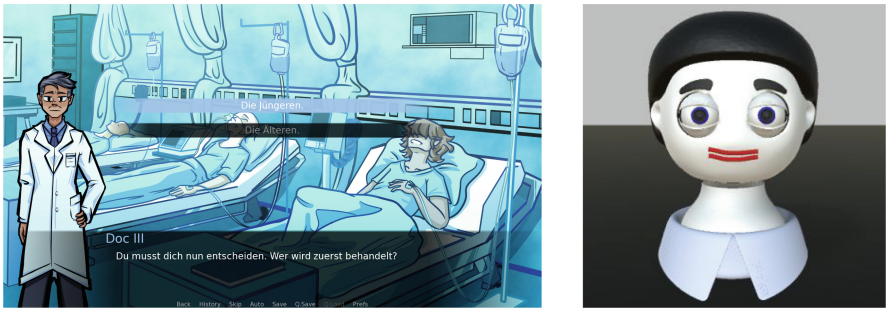


Fig. 1. The study setup: Participants had access to two displays. They interacted with the visual novel on the left display, while seeing Floka on the right one.

3 Method

We designed and created a visual novel with embedded questionnaires to conduct a study. The visual novel introduces the participant to their role as physician in a clinic and leads them through three moral dilemmas: Assisted death for a terminally ill patient, giving an available organ transplant to a stranger or a friend and prioritizing older or younger patients in a triage scenario. The study stays close to [10] with a few notable key differences: (1) It stays in the health sector; (2) it uses an actual machine agent to give advice; (3) the participant first decides without prior advice; (4) it gives contrary advice to the prior decision and prompts the participant if they would change their decision; (5) it measures the participant's perception of the agent via Godspeed questionnaire [3] and their personality via a shortened Big Five [13].

We chose only virtual agent Floka, an improved version of Flobi [12], as machine agent advisor to ensure comparability between its different presentations. The presentation entails Floka telling the participant either nothing about their training data (*neutral*), or revealing the source to be law- and guideline abiding expert opinions (*trustworthy*) or lay people with a criminal record (*untrustworthy*). The presentations follow the advisors in [10]. Any one participant only ever sees one of the three presentations. The visual novel conveys the hospital setting to the participant through text and illustration. As an AI advisor, Floka is not only seen, but also heard, as Floka relays advice through generated speech.

4 Results

Participants fall into three groups according to the agent's presentation: untrustworthy, neutral and trustworthy. Each group has 10 participants, totaling 30 participants in the entire

Table 1. Amount of changed decision after advice by presentation

Advice Taken	2x	1x	0x
Trustworthy	1	6	3
Neutral	1	4	5
Untrustworthy	2	3	5

study. The participants are not representative of the overall population, since most of them are students from Bielefeld University. However, we found no significant differences in personality between groups. All significance tests according to the three groups are determined by the Kruskal-Wallis test. While Advice was only followed in 21 out of 90 decisions, taking said advice shows no dependency on presentation (Table 1 or personality. It is of note that none of the 30 participants changed their decision in the euthanasia scenario, no matter their prior decision. We couldn't determine any significant differences between the entirety of Godspeed items and groups. It follows that the perception of Floka did not significantly vary between presentations (Fig. 2). Hence we further looked into differences in advice taking behavior regardless of presentation groups Fig. 3.

Correlations between amount of advice taken across personality and Godspeed results as indicated by the Pearson test showed positive correlation for *Likability* ($r(28) = 0.49$, $p = 0.006$) and *Intelligence* ($r(28) = 0.44$, $p = 0.015$), see Fig. 4. While the presentation appears to have no influence on the advice taking, the participant's perception of the agent's *Likability* and *Intelligence* does. What influences this perception is unclear, however.

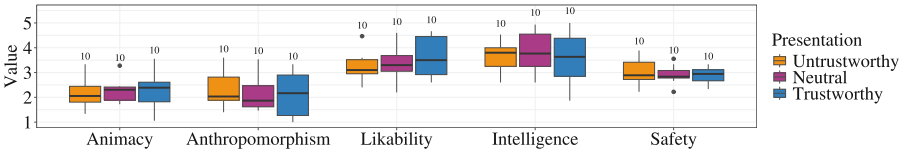


Fig. 2. Perception of Floka measured via Godspeed, grouped by presentation

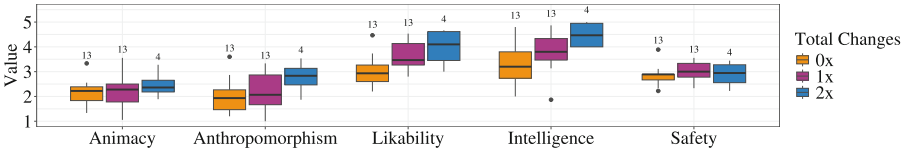


Fig. 3. Perception of Floka measured via Godspeed, grouped by advice taking behavior of participants (amount of changed decisions)

5 Discussion and Future Work

Our study design choices that further focus the AI compared to Krügel et al. still resulted in similar data. The design includes increased immersion via video game interface (visual novel) and an embodied virtual agent capable of speech to cover a full audiovisual range of engagement. We found that, even under the premise

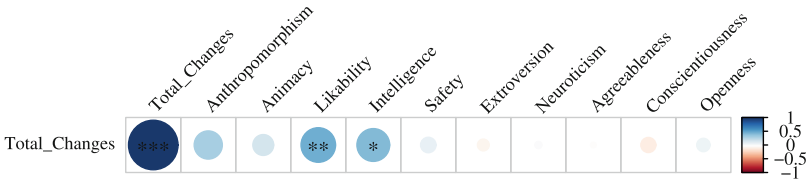


Fig. 4. Pearson Correlation between Godspeed, Big5 items and amount of changed decisions

above, training data transparency implicating a grade of trustworthiness does not meaningfully impact trust in AI. Additionally, indicated untrustworthiness doesn't significantly affect perception of AI measured via Godspeed. Not even on *Safety*, where a priori, at least a small correlation could reasonably be expected (compare criminal human advisors [10]). Hence transparency of training data may not have the effect desired by [7] for end-users.

Further, we found significant correlation between the amount of changed decisions per participant and ratings of *Likability* and *Intelligence*. However, there is no clear indication which factors influence this perception. Neither personality of the participant nor presentation of the agent shows correlation.

It is left to future research to further inquire what influences the advice taking behavior and whether transparency of training data serves as a tool for critical assessment of AI-agents for different stakeholders or in different situations.

Acknowledgments. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): TRR 318/1 2021-438445824 "Constructing Explainability".

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alarcon, G.M., Capiola, A., Pfahler, M.D.: Chapter 7 - the role of human personality on trust in human-robot interaction. In: Nam, C.S., Lyons, J.B. (eds.) Trust in Human-Robot Interaction, pp. 159–178. Academic Press (2021)
2. Bahner, J.E., Elepfandt, M.F., Manzey, D.: Misuse of diagnostic aids in process control: the effects of automation misses on complacency and automation bias. Proc. Human Factors Ergonom. Society Annual Meet. **52**(19), 1330–1334 (2008). https://doi.org/10.1177/154193120805201906
3. Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int. J. Social Robot. (2009)
4. Danks, D.: Governance via explainability. In: Oxford Handbook of the Governance of AI. Oxford University Press (2022)
5. Ferrario, A., Loi, M.: How explainability contributes to trust in AI. In: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, pp. 1457–1466 (2022)

6. Goddard, K., Roudsari, A., Wyatt, J.C.: Automation bias: a systematic review of frequency, effect mediators, and mitigators. J. Am. Med. Inform. Assoc. **19**(1), 121–127 (06 2011). https://doi.org/10.1136/amiajnl-2011-000089
7. HLEG-AI, FUTURIUM, Commision, E.: Ethics guidelines for trustworthy AI (2023)
8. Holzinger, A., Langs, G., Denk, H., Zatloukal, K., Müller, H.: Causability and explainability of artificial intelligence in medicine. WIREs Data Min. Knowl. Discov. **9**(4), e1312 (2019)
9. Kim, S.S., Watkins, E.A., Russakovsky, O., Fong, R., Monroy-Hernández, A.: " help me help the ai": Understanding how explainability can support human-ai interaction. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1–17 (2023)
10. Krügel, S., Ostermaier, A., Uhl, M.: Zombies in the loop? humans trust untrustworthy ai-advisors for ethical decisions. Philosophy & Technology (2022)
11. Lyons, J.B., Nam, C.S., Jessup, S.A., Vo, T.Q., Wynne, K.T.: The role of individual differences as predictors of trust in autonomous security robots. In: 2020 IEEE International Conference on Human-Machine Systems (ICHMS), pp. 1–5 (2020)
12. Lütkebohle, I., et al.: The bielefeld anthropomorphic robot head "flobi". In: 2010 IEEE International Conference on Robotics and Automation, pp. 3384–3391 (2010)
13. Rammstedt, B., Kemper, C.J., Klein, M.C., Beierlein, C., Kovaleva, A.: Big five inventory (bfi-10). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS) (2014)
14. Robinette, P., Li, W., Allen, R., Howard, A.M., Wagner, A.R.: Overtrust of robots in emergency evacuation scenarios. In: 2016 11th ACM/IEEE international conference on human-robot interaction (HRI), pp. 101–108. IEEE (2016)
15. Smuha, N.A.: The EU approach to ethics guidelines for trustworthy artificial intelligence. Comput. Law Rev. Int. **20**(4), 97–106 (2019)

Designing Buy-In: Leveraging HCI to Drive Responsible AI

Yomna Elsayed(✉)

Credo AI, 4546 El Camino Real B10 #795, Los Altos, CA 94022, USA
yomna@credo.ai

Abstract. AI governance establishes organization-wide policies to regulate AI development and use while mitigating risks. Traditionally, it focused on technical teams ensuring transparency in machine learning (ML) models. However, the rise of Large Language Models (LLMs) has expanded AI governance into an enterprise-wide initiative. Generative AI has democratized access to AI, but many users lack training to assess outputs critically. This shift requires a socio-technical approach to governance that spans technical and non-technical teams across HR, sales & marketing, and operations. AI governance professionals face challenges like stakeholder resistance and lack of executive sponsorship, making organizational Buy-in crucial. Buy-in operates at two levels: top-down and bottom-up. Executive support legitimizes governance efforts, but mandates alone are ineffective without grassroots adoption. Conversely, grassroots engagement requires education, collaboration, and trust-building. Drawing from Participatory Design (PD) and Socio-Technical Systems (STS) Theory, this research presents an HCI-informed approach to embedding governance in workflows. While AI ethics and technical tools are widely studied, few address sustainability of governance adoption. This paper provides actionable, HCI-inspired strategies to design governance processes that are user-friendly, seamlessly integrated, and encourage responsible AI adoption across diverse teams.

Keywords: AI Governance · HCI · Participatory Design

1 Introduction

AI governance is the establishment of organization-wide frameworks and policies that regulate the development, deployment and use of Artificial Intelligence within organizations to reduce its operational, regulatory, privacy, security and reputational risks. Prior to the introduction of large language models, AI governance was largely confined to the technical realm of model developers to ensure that traditional machine learning models (ML) are not biased and provide transparent outcomes. However, with the proliferation of Large Language Models (LLMs), aka. Generative AI, AI governance has now become an enterprise-wide initiative designed to de-risk AI and help companies build both internal and external trust in their solutions.

Generative AI has democratized access to AI, albeit without the necessary or proper training to use it responsibly. Many new users lack training to effectively prompt or

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 269–279, 2025.
https://doi.org/10.1007/978-3-031-94171-9_25

scrutinize outputs, leading to over-reliance and misplaced trust in the technology. This shift requires AI governance professionals to develop socio-technical frameworks that cater to the needs of both data scientists and non-technical teams, including sales, marketing, HR, and operations. Additionally, it calls for a broader more integrated approach that encompasses the whole organization in embedding Responsible AI (RAI) practices within organizations.

This research is grounded in professional experience and draws on generalized observations of enterprise AI governance practices. In-depth interviews and design thinking workshops with AI governance professionals reveal that AI governance professionals face many challenges in establishing governance frameworks and processes from grassroots resistance to lack of executive sponsorship. Therefore, securing Buy-in (agreement with, or acceptance of, a policy or suggestion) is crucial to the success of governance processes and their organizational-wide adoption.

Buy-in operates at two levels: top-down and bottom-up. Executive sponsorship formalizes and prioritizes governance policies, serving as the "mandate" that drives compliance. However, top-down mandates have proven ineffective without grassroots support and adoption [8, 10]. At the same time, grassroots Buy-in requires careful cultivation—AI governance professionals must engage teams, provide educational resources, and design tools that facilitate collaboration and build trust over time.

By drawing on principles of Participatory Design (PD) and Socio-Technical Systems (STS) Theory, this paper presents an approach to AI governance informed by Human-Computer Interaction (HCI), which examines how people engage with technology to improve usability and adoption. Previous research has focused on identifying guidelines and principles for the ethical development of AI, and the associated technical tooling. However, limited research explored the biggest challenge currently facing AI governance professionals, which is how to embed governance in organizational processes by securing long-lasting Buy-in with various stakeholders.

This paper provides critical guidance on how to build governance tools and processes that are widely adopted by diverse users with distinct needs. By applying user-centered principles, governance professionals can bridge both technical and non-technical spaces, create seamless processes, and build awareness and trust in RAI practices. While executive sponsorship is crucial for ensuring AI governance is prioritized, this paper will provide HCI-inspired steps that governance professionals can follow to design grassroots compliance through accessible, actionable frameworks that empower teams to adopt AI responsibly.

2 Related Work

AI Governance is seldom approached as an HCI problem, and is usually treated as a technical problem that emphasizes bias and fairness testing. [1] proposed a holistic AI governance framework (HAGF) aimed at addressing bias and ensuring accountability through a focus on transparency and continuous monitoring. While the HAIG framework offers valuable insights into high-level governance techniques, it primarily focuses on organizational policies and oversight mechanisms, leaving out the socio-technical aspects of actually integrating that framework into the daily workflows of AI practitioners. As [4] points out, AI governance is a socio-technical problem that involves social

factors such as stakeholders, processes, and cultures alongside technical aspects. [12] emphasizes a human-centric approach to AI governance, arguing that existing policy discourses often fail to fully operationalize human-centered AI (HCAI) principles in governance practices. Their work advocates for a systemic approach that incorporates transparency, inclusiveness, and participatory decision-making. However, it frames governance in the public governance and policy domains which are distinctly different from corporate settings, where governance teams have the extra challenge of showing ROI for their governance processes.

Participatory design principles have been referenced in a couple of previous works as guiding principles for governance frameworks. For Example, [6] notes how adopting flexible and diverse governance structures that involve multiple stakeholders enables organizations to better manage the various risks associated with AI technologies. [15] similarly emphasizes that the lack of grassroots involvement and coordination between top-down and bottom-up approaches to governance can lead to policies that are disconnected from practical needs, thus emphasizing the importance of participatory practices in governance. However, few works explicitly explore governance frameworks inspired by participatory design principles.

Due to the nascent nature of the AI governance field and the lack of clarity around stakeholder roles, little research has addressed the organizational deployment of AI governance processes. [2] Therefore advocate for more research to clarify organizational AI governance processes and actor roles. This paper introduces a crucial socio-technical perspective that views AI governance as an organizational challenge shaped by technology. Without well-defined processes and participatory mechanisms, these technologies will struggle to gain user trust and adoption.

3 Theoretical Framework

Participatory Design (PD) is a collaborative approach that involves end users in the creation process of a service, product or a system. PD highlights the importance of diversifying input and embedding design in community needs through the following key principles:

- Co-creation and user-involvement: co-creation actively involves diverse stakeholders in design discussions to understand their context and pain points, thus ensuring practical-relevance and real world impact.
- Collaborative Learning: Design thinking relies on a shared, social understanding with the stakeholders who will later use the tools [7]. Therefore, collaborative learning ensures the continuous exchange of knowledge between different stakeholders, and adoption to evolving technologies.
- Empowerment: Empowerment is about making sure the users of the system have necessary knowledge and tools to take ownership of the systems and the decisions that shape it [14].
- Usability and transparency: Usability and transparency are about making sure that systems are intuitive, accessible and aligned with user needs [9], fostering trust, reliability and wide-spread adoption among users.

Socio-Technical Systems (STS) Theory is another theoretical framework suitable for the analysis of AI governance interaction within organizations due to its emphasis on the interconnectedness of people, policies and technologies. Ultimately, AI governance tools and processes have to be embedded within existing organizational workflows to ensure sustainability and long term adoption. Process alignment, distributed accountability and responsible autonomy are key STS principles that AI governance teams can draw upon to secure Buy-in for their processes.

- Process alignment: Is the act of reconciling or jointly optimizing [3] social and technical elements within an organization for harmonious integration.
- Distributed Accountability: Is about shared responsibility, where accountability does not rest solely on one person's shoulders, but is rather distributed through clear roles and oversight mechanisms.
- Responsible Autonomy: happens when stakeholders take ownership of the entire process, while also acknowledging the interdependence between individuals or groups to ensure a smooth workflow [13].

By leveraging both PD and STS principles as a foundation for building AI governance processes and securing Buy-in for it, AI governance teams can design processes that are not only effective but also sustainable.

4 Research Methodology

4.1 Study Design

This study draws on in-depth interviews and design thinking workshops with various stakeholders involved in the AI governance process, primarily AI governance professionals, but also data scientists, legal and compliance, IT, security and privacy officers as well as executive leadership. This diverse representation provides a holistic perspective on governance that spans the entire organization.

The study was conducted between 2023 - 2024 and hence captures the changing attitudes towards AI and governance coincident with the introduction of generative AI. These interviews and workshops were conducted within the author's professional environment allowing first-hand accounts from industry practitioners and ensuring the findings are deeply rooted in real-world practices with strong practical relevance.

4.2 Data Collection and Analysis

Interviews and workshops focused on AI governance challenges to adoption and Buy-in strategies. Discussions also examined governance processes, change management and risk mitigation strategies within AI-driven organizations. A thematic analysis was performed to extract key insights from qualitative findings identifying key patterns related to adoption challenges, and Buy-in strategies.

4.3 Limitations and Scope

This study is focused on private enterprise settings and may not be generalizable to public sector AI governance or regulation. Additionally, while the study provides insights from industry professionals, its findings may not be applicable across all organizational structures particularly non-corporate or educational settings.

5 Findings and Insights

5.1 Challenges to Buy-In an Adoption

Governance processes are often poorly adopted due to challenges in institutionalizing and integrating them into existing organizational frameworks. A recurring pain point is stakeholder's resistance to governance. Governance is widely seen as a barrier rather than a catalyst for innovation and operational efficiency. However, resistance to governance may be less about opposition and more about lack of empowerment, support structures, and alignment with existing development processes. Participants highlighted the following reasons as key contributors to this resistance:

Lack of Usability. Governance workflows and tools are often too manual and complex for seamless adoption. AI governance professionals either rely on a variety of general purpose tools to perform governance such as spreadsheets and surveying tools, or they purchase AI governance-specialized solutions designed to automate data collection and review for AI. However, many of these AI governance tools struggle to keep pace with the disruptive evolution of AI technologies. Rather than offering cohesive seamless experiences, they frequently rely on disjointed add-ons that attempt to superficially address emerging AI risks. This reactive approach neglects the institutional dimension of AI governance, failing to embed governance processes into broader organizational frameworks in a scalable and sustainable way.

Stakeholders are looking for a governance experience that minimizes effort, focuses attention and is actionable. Therefore, participants expressed a need for a "zero-clicks" and "plug-and-play" experience. They also emphasized the importance of having a natural workflow that minimizes the number of steps it takes to advance AI use cases confidently to deployment.

Lack of Awareness and Expertise. Even with a seamless experience, AI and RAI in specific remain to be new and challenging areas that few people have the knowledge and expertise to understand. AI governance professionals themselves are continually learning and adapting their processes to AI advancements.

According to [11], few computer science courses address the ethical implications of AI which contributes to a crisis in implementing ethical AI. Few course syllabi mention ethics-related topics (privacy, bias, fairness) [6]. Therefore, while machine learning model developers might be very adept at calculating metrics such as precision, recall and performance, they may not be equally equipped to calculate metrics such as fairness and bias. Lack of access to representative data, fairness-accuracy tradeoffs, and lack of tooling are key contributors to lack of governance adoption.

Additionally, AI governance administrators who ask technical stakeholders to calculate RAI metrics, may lack the knowledge and training required to interpret and act upon such metrics. While some AI governance professionals are data scientists who turned to AI governance out of concern for AI risks, others come from legal, compliance, IT backgrounds and may thus lack the technical ability to define the acceptable thresholds or interpret the outcome of these RAI metrics.

Lack of Alignment. AI governance define processes for stakeholders to disclose AI use, assess its risk, and implement mitigations for it. These processes, however, add complexity when they are not aligned with existing AI development processes. Stakeholders often struggle with when and how to seek governance, especially when AI can be embedded into newly purchased tools or legacy software looking to enhance their tools with AI. This expands the purview of AI governance from 1st party (internally) developed models and tools to the expansive space of 3rd party tools, and risk. Hence, many AI governance professionals find themselves in the position of having to reconcile their processes with both the procurement and model development life cycles.

Reconciliation even extends to existing AI governance processes. Oftentimes, AI governance professionals experience great pains in fitting their existing manual governance processes with the automated workflows dictated by a newly-purchased governance tool, and later convincing their stakeholders to adopt a newer process. This can sometimes result in a tradeoff between Buy-in and efficiency.

AI developers are responsible for ML monitoring processes embedded in the AI development lifecycle and CI/CD pipelines, tracking key performance indicators (KPIs) like accuracy, precision, recall, drift detection, and retraining needs. An additional AI governance layer may be seen as disruptive, introducing compliance checks, documentation requirements, and validation steps that could slow model iteration cycles and impact business agility.

Lack of Incentive. Executive leadership's lack awareness of AI risks and their broad operational, legal, and reputational implications may amplify the above challenges. Without a clear understanding of AI risk, leaders may struggle to implement effective governance strategies, potentially leading to compliance failures, ethical missteps, and an erosion of trust between the organization and its clients. Additionally, the absence of regulatory urgency – due to bureaucratic red tape, slow moving policy developments or lack of enforcement– can allow organizations to deprioritize AI governance. This regulatory lag can create an environment where risk management is more reactive than proactive subjecting companies to legal liabilities and reputational damage.

Without executive Buy-in and sponsorship, AI governance professionals face an uphill battle of stakeholder education on both grassroots and leadership fronts. The challenge is further compounded by lack of funding which can lead to inadequate compliance efforts and inefficient governance processes. Fortunately, generative AI is driving change, as its high risks and business value make governance concerns impossible to ignore. Many executives now recognize that innovation with generative AI requires robust guardrails, ultimately linking governance to tangible ROI.

5.2 HCI-Informed Strategies for Establishing Buy-In

In addition to sharing their governance challenges, participants shared their strategies for overcoming organizational resistance to governance and securing Buy-in. Many of their approaches fit well with concepts of participatory design and socio-technical theory in their emphasis on human-centered approaches to the design and implementation of governance processes:

Multi-Stakeholder Participatory Design of Governance Processes. Participatory design principles emphasize the importance of co-design and co-creation. Co-design is the process of involving multiple stakeholders in the design and development of services, products or systems. By its nature, co-design is a collaborative process that aims to address users' needs and achieve their objectives. This way designers can pave the way for increased user satisfaction and engagement.

Inclusion. GenAI democratized access to AI, making it accessible to entire departments that have limited to no training in AI, such as sales, HR and marketing. To address GenAI risk, AI governance teams across enterprises formed enterprise-wide multi-stakeholder committees that spanned privacy, security, legal, DEI, and executive leadership to address the risks of GenAI. As STS emphasizes the role of strong leadership in facilitating collaboration and embedding processes in user needs [5], executive leadership act as sponsors to theses committee's efforts by making sure their policies are organizationally embedded. Multi-stakeholder governance committees facilitate inclusive decision making, central to principles of participatory design.

Co-Creation. AI committee stakeholders collaborate to develop a structured set of questions for evaluating new AI use cases, ensuring relevance to their areas of expertise while avoiding redundancy. They also establish policies that address both domain-specific and organization-wide concerns, aligning with participatory design's emphasis on embedding solutions within community needs. Committees meet regularly to review and approve AI use cases in a transparent manner. Transparency fosters awareness over time. This awareness could eventually evolve into self-governance of AI use cases, where Buy-in is driven by an internal socio-technical commitment to building trustworthy and responsible AI products. Self-governance is a reflection of STS's emphasis on "responsible autonomy," whereby teams are empowered to make governance-aligned decisions without excessive bureaucracy [13].

Embedding Governance into Existing Workflows. To address the misalignment between governance and existing processes (e.g. development life cycle), governance admins iteratively work with stakeholders to align key stages and steps. According to STS (Socio-technical theory), the technical and social dimensions of AI development should co-evolve in harmony to ensure balanced and sustainable processes. Some of the ways that governance teams went about this are:

Workflow Customization. By dynamically adapting intake questionnaires based on the relevant phase of the development process, governance admins ensure that different phases address distinct governance considerations. Participants also highlighted the

value of creating separate Key Performance Indicators (KPIs) and guidelines for approving different stages of the AI life cycle to ensure that governance is tailored to the unique risks, challenges, and requirements at each phase.

Alignment with Existing Processes. Governance teams ensure that ethical, security, and compliance considerations are embedded seamlessly into the AI lifecycle, by closely studying the CI/CD (Continuous Integration/Continuous Deployment) pipeline to align AI governance with the existing development workflows. By including an important governance step at each milestone of the development life cycle, governance teams embed governance in their technical stakeholders' day-to-day operations all while gradually building governance awareness.

Technical Integrations. Participants emphasized the importance of technical integrations with stakeholders' platforms and tools to minimize friction and avoid duplication of work. They integrate workflow and ticketing systems, privacy and security platforms, and MLOps tools to facilitate the flow of information, making AI governance the centralized pane from which stakeholders could monitor compliance, track approvals, and manage risks, without disrupting existing workflows.

Transparent and Usable Governance Tools. Governance processes need to be frictionless to gain stakeholders' adoption. While some stakeholders are willing to participate in governance out of genuine concern about its risk, they may be discouraged by logging into multiple platforms, filling out extensive questionnaires, or going through a long approval cycle. Therefore, governance teams put usability on top of their process requirements. According to PD, usability enables users to experience success with AI governance, hence ensuring process sustainability.

Transparency is also crucial in PD for fostering trust in the process. Governance teams maintain transparency through governance tools with communication capabilities that help them provide necessary context and ask questions about governance. Governance teams make sure that their stakeholders are well aware of the governance timeline, stage they are in, and what to expect at each stage. Throughout the AI risk assessment, users are included in iterative feedback loops where they are not only supplying information but also being educated about context-specific AI risk.

Empowerment Through Technical Tooling and Dashboards

Tooling. Cultural controls alone are insufficient for securing buy-in; users must feel empowered to engage in AI governance. A major challenge is that even computer scientists often lack training in RAI and its assessments [6, 11]. Hence, governance admins bridge this gap, which can be overwhelming. Governance tooling becomes essential through the use of RAI experts, SaaS solutions, or consulting services.

These experts engage with stakeholders to understand their knowledge gaps and work with them on providing reusable governance solutions. Collaborative engagements empower users to adopt governance and be less resistant to change. Bias and fairness testing APIs are one way where participants attempt to reduce the knowledge gap and streamline their processes. More tech-savvy participants have also indicated that they

provide data science teams with mechanisms to assess fairness using their own proprietary bias evaluation methods and datasets, all meant to empower data science teams with the necessary tools to unblock governance.

Dashboards. Another way participants keep their stakeholders engaged is through dashboards that enable stakeholders to stay updated with governance insights. They can also educate stakeholders about their AI risk profile, track their submissions, and how they are complying with regulations. In addition to offering transparency into a seemingly complex process, these dashboards can also offer insights about the bias and fairness of their AI. By making governance more transparent, interactive, and actionable, these dashboards align with PD principles of co-creation and sustained user engagement, essential for a sustainable and scalable governance process.

Collaborative Learning. Getting Buy-in would not be possible without the necessary ground work of building awareness as to the risks of AI. According to [2], staff training and competencies are crucial for AI governance adoption. Awareness is also essential for securing executive sponsorship, especially when leadership is eager to innovate with GenAI without weighing its risk. Fortunately, this awareness can develop naturally as public discourse highlights cases of GenAI misuse with repercussions, such as biased hiring models, hallucinations, IP violations, and security breaches.

While participants embed AI governance training in company-wide training materials, they generally find hands-on training—achieved through process adoption and engagement—to be more valuable. One participant noted that it is important for model developers, not governance admins, to run their own bias test, so they can continue to build their understanding and maintain accountability for their AI projects. Obviously, this is not a straightforward endeavor and usually involves a lot of setbacks and iterations for AI governance professionals to build awareness about AI risks.

6 HCI-Informed Framework for AI Governance

To build a governance process that is resilient and adaptive to both organizational and technological change, governance teams need to center these guiding principles:

6.1 Guiding Principles

- Human Centeredness: We build processes that adapt to human needs to ensure practicality and wide-adoption.
- Diversity & Inclusion: We invite a wide-range of stakeholders to ensure inclusivity and real-world applicability.
- Co-creation: We actively involve stakeholders in the design and implementation of governance frameworks to ensure governance is iterative and context-aware.
- Embeddedness: We integrate seamlessly with existing workflows rather than disrupting them, ensuring governance efficiency and smooth adoption.
- Transparency: We communicate governance processes, decisions, and evaluation criteria clearly to stakeholders to ensure governance accessibility and trust.

- Empowerment: We empower stakeholders with usable tooling to help them meet governance requirements and encourage ownership and accountability.
- Collaborative Learning: We create environments where stakeholders share knowledge and feedback ensuring continuous improvement of the process.

6.2 HCI-Inspired Governance Framework

An HCI-inspired governance framework takes into account all the above principles, all while making sure to introduce them at the most appropriate time:

- Step 1: Conduct participatory workshops: Invite stakeholders from security, privacy, legal, data science, DEI, and executive leadership to recurring participatory workshops to understand specific contexts, and identify user pain points and needs.
- Step 2: Create a collaborative learning environment: Hold monthly Responsible AI share-outs, to exchange knowledge on AI risks and best practices inviting experts to weigh in from their different areas of expertise.
- Step 3: Co-design governance processes: Conduct design thinking sessions to map out existing workflows, identify gaps and co-develop processes.
- Step 4: Implement governance in phases with feedback loops: Rollout the governance process in phases that are aligned with teams' evolving maturities, introducing more complexity as they mature. Collect feedback regularly to refine practices.
- Step 5: Co-design dashboards for transparency: Once the governance process is established, conduct participatory workshops to identify what insights, metrics, and compliance details each team needs to see aggregated, ensuring dashboards provide clear, actionable, and team-specific insights about the governance process.
- Step 6: Empower teams with tooling and integrations: As teams mature in disclosing and assessing AI use cases, identify governance tooling needs, automate checks, and streamline communication—ensuring governance is continuous, not one-time.

7 Conclusion

GenAI has pushed AI governance from the technical realm into the socio-technical realm, whereby governance technologies need to be embedded in and conversant with their socio/organizational context. Using HCI inspired approaches that draw upon Participatory Design Principles and Socio-Technical Theory, this paper presented a 6-step framework for securing Buy-in that is informed by real-world challenges and needs from industry AI governance practitioners. To ensure relevance, adoption and sustainability of the governance process, this framework is rooted in the principles of human-centeredness, diversity, co-creation, embeddedness, transparency, empowerment and collaborative learning. This work was largely qualitative in nature relying on long-form interviews and design thinking workshops. Future work could explore quantitative adoption metrics, essential for demonstrating the ROI of governance and its positive impact on innovation through effective risk management.

Acknowledgments. This work was made possible with the support of Credo AI, which provided resources, settings and access to participants. The views expressed in this paper are my own and do not necessarily reflect those of Credo AI.

References

1. Atoum, I.: Revolutionizing AI Governance: Addressing Bias and Ensuring Accountability Through the Holistic AI Governance Framework
2. Birkstedt, T., Minkkinen, M., Tandon, A., Mäntymäki, M.: AI governance: themes, knowledge gaps, and future agendas. Internet Res. **33**(7), 133–167 (2023)
3. Di Maio, P.: Towards a metamodel to support the joint optimization of socio-technical systems. Systems **2**(3), 273–296 (2014)
4. Feng, K., et al.: Sociotechnical AI Governance: Challenges and Opportunities for HCI. In: CHI 2025 Workshop on Sociotechnical AI Governance (2025)
5. Hazy, J.K.: Measuring leadership effectiveness in complex socio-technical systems. Emergence: Complex. Organizat. **8**(3) (2006)
6. Garrett, N., Beard, N., Fiesler, C.: More than "If Time Allows": the role of ethics in AI education. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, pp. 272–278 (2020)
7. Häikiö, E., Heikkilä, J.: Adoption and Governance of AI-Powered Dashboards in Executive-Level Decision-Making. Master's Thesis
8. Lupton, T.: Organisational change : "top-down" or "bottom-up" management ? Pers. Rev. **20**(3), 4–10 (1991)
9. Pope, A.R.: Usability, Use, and Participatory Design. Open Technical Writing. https://tinyurl.com/4bc6dwcw
10. Ryan, N., Williams, T., Charles, M., Waterhouse, J.: Top-down organizational change in an Australian Government agency. Int. J. Public Sect. Manag. **21**(1), 26–44 (2008)
11. Raji, I.D., Scheuerman, M.K., Amironesei, R.: You Can't sit with us: exclusionary pedagogy in AI Ethics education. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 515–525 (2021)
12. Sigfrids, A., Leikas, J., Salo-Pöntinen, H., Koskimies, E.: Human-centricity in AI governance: a systemic approach. Front. Artifi. Intell. **6** (2023)
13. Trist, E.L., Higgins, G.W., Murray, H., Pollock, A.B.: Organizational choice. capabilities of groups at the coal face under changing technologies. the loss, re-discovery and transformation of a work tradition. Routledge, London (2013 (1963))
14. Zhang, M.: Design Empowerment: Participatory Design Towards Social Sustainability. In: Rau, PL.P. (eds.) Cross-Cultural Design. Interaction Design Across Cultures. HCII 2022. LNCS, vol. 13311. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-06038-0_20
15. Zhong, H., Do, T., Jie, Y., Neuwirth, R.J., Shen, H.: Global AI Governance: Where the Challenge is the Solution—An Interdisciplinary, Multilateral, and Vertically Coordinated Approach. arXiv preprint arXiv:2503.04766 (2025). 15

How Perceived Competence of Generative AI Affects Self Efficacy

Ron Fulbright(✉)

University of South Carolina Upstate, 800 University Way, Spartanburg, SC 29303, USA
fulbrigh@uscupstate.edu

Abstract. Self-efficacy refers to an individual's belief in their own ability to successfully achieve a specific task or goal. A person with higher self-efficacy is likely to be more motivated and willing to take on challenging tasks, persist in the face of adversity, and recover quickly from setbacks. Self-efficacy in the context of artificial intelligence (AISE) refers to how confident users feel about using and adopting AI tools. When an AI tool is perceived as reliable, trustworthy, and widely adopted, users are more likely to feel confident in their ability to use it successfully. This paper explores the effect preconceptions, based solely on the name of the tool, have on people's assessment of the AI tool. We find if an AI tool is widely known or if the name of the AI tool specifically mentions the task, people will assess the tool higher, even if the tool is not better in reality. Therefore, a person's preconceptions affect their self-efficacy.

Keywords: artificial intelligence · human-computer interaction · human cognitive augmentation · human/cog ensembles

1 Introduction

Human performance of any kind is augmented by using tools. Physical performance is enhanced by using simple tools like hammers, shovels, and axes. Likewise, human cognitive performance is augmented by using tools able to process and transform information. For example, unaided, a human might take several minutes to add a column of numbers, and then the sum would need to be checked because of the possibility of error in the calculations. However, using a calculator or a computer spreadsheet, a human could produce a reliable sum in a fraction of a second. In fact, entry of the numbers becomes the limiting factor in terms of speed. Today, we commonly use software to process words, images, video, and numbers in our daily work. Such tools have taken over the burden of lower-level cognitive "grunt work."

Recently, however, cognitive systems technology ("AIs") built using unsupervised, deep, machine learning techniques, has produced tools able to outperform humans in domains formerly thought to require high-level human cognitive processing. When humans use tools like these in a collaborative manner (in a human/AI ensemble) human cognitive performance is enhanced—augmented. If a human's cognitive ability is augmented enough, even a non-expert can achieve, and even exceed, the performance

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 280–284, 2025.
https://doi.org/10.1007/978-3-031-94171-9_26

of an expert in a particular domain—something called synthetic expertise (Fulbright 2016a; 2016b; 2017; 2018; 2019; 2020; 2020a; Fulbright & Walters 2020, Fulbright & McGaha 2023; Fulbright & Morrison, 2024).

Systems like the Chat Generative Pre-Trained Transformer (ChatGPT), have gained much attention recently. ChatGPT is a large language model trained to predict the most probable next word in a sequence of words and is fine-tuned for conversational usage. ChatGPT mimics human-created text. Instead of being limited to a narrow domain, users can conduct extended textual dialogs with ChatGPT on practically any topic and most of the time, text generated by ChatGPT is indistinguishable from text produced by another human. Every day, millions of people use ChatGPT for assistance in learning, researching, getting advice, writing music, poetry, and prose, generating computer program code, and much more.

However, the level of cognitive augmentation depends on how well the human uses the AI tool. Important here is how the human feels about the AI tool in the first place. Self-efficacy theory (Bandura 1977) refers to an individual's belief in their own ability to successfully perform and achieve specific tasks or goals. A person with higher self-efficacy is likely to be more motivated and willing to take on challenging tasks, persist in the face of adversity, and recover quickly from setbacks. Self-efficacy in the context of artificial intelligence (AISE) refers to how confident users feel about using and adopting AI tools (Shneiderman 2022). When an AI tool is reliable, trustworthy, and widely adopted, users are more likely to feel confident in their ability to use it successfully (Riedl 2022).

Wang and Chuang (2024) propose an instrument to measure an individual's Artificial Intelligence Self-Efficacy (AISE). The instrument seeks to gauge users' rating of an AI tool's understandability, comfort, ease of use, and helpfulness. However, none of the 22 items in the AISE instrument asks about the user's pre-conceived perception or previous knowledge of the AI tool being evaluated. We question if previous knowledge or perceived competence of an AI tool affects a person's rating of an AI tool. The hypothesis investigated in this paper is:

> **H1:** People will assess an AI tool's performance higher if they have previous knowledge of the AI tool or perceive the tool as being specifically intended for the task at hand.

2 The Experiment

To test hypothesis H1, we designed an experiment asking respondents to rate the output of a letter generated by the Generative AI ChatGPT-4 resulting from the prompt:

> *Write a heartfelt one-paragraph letter to my boss saying I will be leaving my job after two more weeks to care for my aging mother who recently experienced a devastating stroke and now needs continuous care.*

All respondents were given the same letter generated by ChatGPT. One-third of respondents were truthfully told the letter was generated by ChatGPT. One-third of respondents were told the letter was generated by a fictitious AI tool called "LetterWriter

3000" and one-third of respondents were told the letter was generated by a fictitious AI tool called "X10." LetterWriter 3000 was used as the name because it sounds as if it was created specifically for writing letters of the type requested in the prompt. X10 was used because of its generic nature carrying no implied connection to the task of letter writing. Of course, ChatGPT is well-known and even if respondents had not used it themselves all respondents had heard of ChatGPT and had heard of fantastic results obtained from using ChatGPT.

Respondents were asked to score the quality, word usage, and tone of the letter on a 0–10 point scale (0 being the worst and 10 being the best). Since the same letter was presented to all respondents, any difference in scoring can be attributed to the respondent's pre-conceived expectations simply based on the name of the AI tool.

Prior to the experiment, we expected ChatGPT respondents to rate the letter the highest because of the world-wide reputation of ChatGPT. We wondered if respondents would rate LetterWriter 3000 higher than X10. We theorized LetterWriter 3000 would be rated higher than X10 because of the specific-sounding name. It sounds like the tool was designed specifically for the task at hand (writing letters). We also expected both Letter Writer 3000 and X10 output would be rated lower than ChatGPT output.

3 Results and Conclusion

Results verified our hypothesis. ChatGPT was rated the highest in all three evaluation categories. X10 output was rated the lowest in all three categories. LetterWriter 3000 output was rated in the middle, but very close to ChatGPT shown in Table 1 and Fig. 1.

Table 1. Rating of AI Tools

Name of Tool	Quality	Words	Tone	Average
ChatGPT	8.33	.89	8.89	8.37
LetterWriter 3000	8.29	.71	8.57	8.19 (-2.1%)
X10	8.00	8.00	7.67	89 (-5.7%)

Since the letter every respondent evaluated was the same, any difference in the ratings must be due to how the person reacted to the name of the AI tool. Since ChatGPT is so well known, people expect it to produce superior results. Therefore, we feel respondents were predisposed to rate the output higher if they knew it came from ChatGPT. Likewise, since no one knew anything about an AI tool called X10, there was no pre-conceived expectation. The difference between the ChatGPT rating and the AI tool is striking.

Interesting is the rating of LetterWriter 3000. Since this name is fictitious and doesn't really exist, respondents had no prior knowledge of the tool, nor how good it might be at the task. Therefore, one would expect it to be rated the same as X10. However, LetterWriter 3000 was rated higher than X10 and almost as high as ChatGPT. The reason for this high rating must be because respondents were impressed with the name.

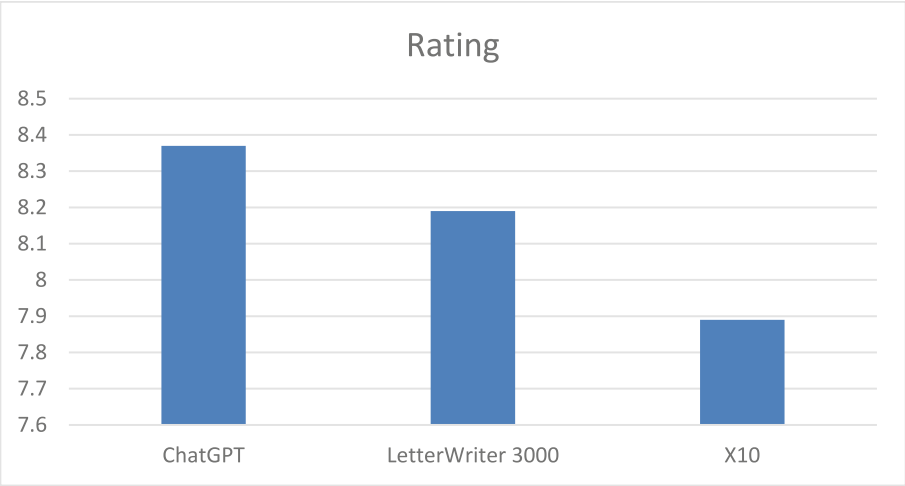


Fig. 1. Rating of AI Tools

LetterWriter 3000 sounds like a tool specifically designed to create excellent letters. We believe this expectation influenced respondents' rating.

A tenet of AI self-efficacy (AISE) is if a person thinks better of an AI tool they will perceive the probability of their own success using the tool to be higher. Our results indicate one's self-efficacy is skewed simply by nominative determinism—the idea people tend to think a name influences the perception or character of the person, place, or thing.

References

Bandura, A.: Self-efficacy: toward a unifying theory of behavioral change. Psychol. Rev. **84**(2), 191–215 (1977)

Fulbright, R.: The Cogs Are Coming: The Cognitive Augmentation Revolution, Proceedings of the Association Supporting Computer Users in Education 2015 (49th, Myrtle Beach, SC) (2016a).. https://eric.ed.gov/?q=cognitive+development+in+early+childhood&ff1=dtysince_2014&pg=1340&id=ED570900, Accessed January 2024

Fulbright, R.: How personal cognitive augmentation will lead to the democratization of expertise. Adv. Cognitive Syst. **4** (2016b). http://www.cogsys.org/posters/2016/poster-2016-3.pdf, Accessed January 2024

Fulbright, R.: Cognitive augmentation metrics using representational information theory. In: Schmorrow, D., Fidopiastis, C. (eds.) Augmented Cognition. Enhancing Cognition and Behavior in Complex Human Environments. AC 2017. LNCS, vol. 10285. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58625-0_3

Fulbright, R.On measuring cognition and cognitive augmentation. In: Yamamoto, S., Mori, H. (eds.) Human Interface and the Management of Information. Information in Applications and Services. HIMI 2018. LNCS, vol. 10905. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-92046-7_41

Fulbright, R.: Calculating Cognitive augmentation – a case study. In: Schmorrow, D., Fidopiastis, C. (eds) Augmented Cognition. HCII 2019. LNCS, vol 11580. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22419-6_38

Fulbright, R.: Democratization of Expertise: How Cognitive Systems Will Revolutionize Your Life. CRC Press, Boca Raton, Fl (2020)

Fulbright, R.: The Expertise Level. In: Schmorrow, D., Fidopiastis, C. (eds.) Augmented Cognition. Human Cognition and Behavior. HCII 2020. LNCS, vol 12197. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50439-7_4

Fulbright, R., Walters, G.: Synthetic Expertise. In: Schmorrow, D., Fidopiastis, C. (eds.) Augmented Cognition. Human Cognition and Behavior. HCII 2020. LNCS, vol. 12197. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50439-7_3

Fulbright, R., McGaha, S.: The Effect of information type on human cognitive augmentation. In: Schmorrow, D.D., Fidopiastis, C.M. (eds.) Augmented Cognition. HCII 2023. LNCS, vol 14019. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35017-7_14

Fulbright, R., Morrison, M. Does Using ChatGPT Result in Human Cognitive Augmentation?. In: Schmorrow, D.D., Fidopiastis, C.M. (eds.) Augmented Cognition. HCII 2024. LNCS, vol 14694 (2024). Springer, Cham. https://doi.org/10.1007/978-3-031-61569-6_9

Riedl, R.: Is trust in artificial intelligence systems related to user personality? review of empirical evidence and future research directions. Electr. Markets **32**, 2021–2051 (2022). https://doi.org/10.1007/s12525-022-00594-4

Shneiderman, B.: Human-Centered AI, Oxford Press (2022)

Wang, Y.-Y., Chuang, Y.-W.: Artificial intelligence self-efficacy: Scale development and validation. Educ. Inf. Technol. **29**, 4785–4808 (2024). https://doi.org/10.1007/s10639-023-12015-w

SJ Urban Impact: Generative AI Powered System for Participatory Civic Engagement Among Young Adults

Michelle Ha and Yu Chen(✉)

School of Information Systems and Technology, San Jose State University, San Jose, USA
{michelle.ha,yu.chen}@sjsu.edu

Abstract. The expansion of capabilities and rising levels of accessibility for AI, especially generative AI, paves the way for innovative solutions to everyday issues. Many young adults experience great difficulty in reaching local authorities to discuss challenges, staying involved in their communities, and having their voices heard. To address this lack of civic engagement in young adults, we created SJ Urban Impact, which is a generative-AI-powered web application to address common barriers and provide a convenient, easily-accessible method to encourage citizens to stay involved in their communities. Through 6 semi-structured interviews with young adult participants living around the Silicon Valley San Jose area, we found that many of them do not participate in their communities because it is difficult to learn about available opportunities and most people do not have the time to get involved. Meanwhile, having a convenient, accessible, and easy-to-navigate intelligent tool could raise motivation for young individuals to stay involved and actively participate in their communities. We discuss the features of the system as well as lessons learned from preliminary UX testing, limitations, and future work.

Keywords: Generative artificial intelligence · Civic engagement · Urban development · Image generation · Text and image analysis · Natural language processing · Community involvement

1 Introduction

The growing number of citizens in urban areas leads to increased problems in infrastructure and location-based problems (Hartmann et al. 2019). Therefore, there is an increasing need for citizens to participate in civic activities with shared governance with the public sectors. Among different age groups, young people are the leaders of the future and thus hold more responsibilities. It is estimated that around 60% of all rural youth live in civic deserts along with 30% of urban and suburban Americans (Attwell et al. 2017). Civic deserts are described as "communities without opportunities for civic engagement" (Attwell et al. 2017). Many barriers contribute to the lack of civic engagement, particularly in young adults, including: difficulty connecting with officials, time constraints due to busy schedules, and a lack of opportunities catered to the specific interests of individuals.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 285–293, 2025.
https://doi.org/10.1007/978-3-031-94171-9_27

Meanwhile, young people are more open to new technologies and thus likely participating in technology-enabled civic engagement. Prior research has shown that technology-enabled participatory platforms for civic engagement can provide significant values for engaging citizens in solving public-good challenges (Desouza and Bhagwatwar 2015). Among them, the recent advancement and adoption of artificial intelligence (AI), especially generative AI, paves the way for innovative solutions to increase citizens' civic engagement. This poster explores the potential of using emerging technologies such as generative artificial intelligence (Gen AI) in reducing barriers to and promoting youth civic engagement.

2 Related Work

Many young adults are utilizing ChatGPT, the most popular commercial Gen AI system, to complete a variety of tasks. Most people use it to aid in productivity-related cases, such as learning more about topics for tasks at work or assignments at school. Furthermore, according to our preliminary study, some use generative AI to discover recommendations for restaurants or sightseeing destinations when they visit new locations. Among the general population of young adults, the four most common uses of generative AI are getting information and brainstorming, followed by assisting with schoolwork and generating images (HopeLab, 2023). Some of the most popular generative AI systems currently used by young users include ChatGPT–an AI chatbot that answers questions and generates text, GitHub Copilot–an AI coding assistant that makes code suggestions to programmers, and DALL-E–a generative AI model used to create images (Sharma 2024).

To leverage these technologies in addressing the lack of civic participation, we explore applying Gen AI to support young adults who want to quickly report city issues, envision their future cities, and catch up on local news with brief summaries. Users can easily interact with our app and generate content in quick iterations. We aim to investigate how Gen AI can empower younger generations to contribute to addressing community issues, especially the city of San Jose in California. Due to the widespread accessibility of text and image generation technologies like the OpenAI API, we based SJ (San Jose) Urban Impact on these capabilities that can provide users with a convenient interface to address community problems.

3 SJ Urban Impact

To investigate how to leverage generative AI to support young people in participatory civic engagement, we prototyped **SJ Urban Impact**, a web application to address common barriers and provide a convenient and easily-accessible method to encourage citizens to stay involved in their communities.

3.1 Scenario

Consider Aaron, a college student living in a shared apartment in downtown San Jose. As someone who aspires to be a more active and engaged member of his community,

Aaron stays involved in local issues through social media. He regularly attends city council and town hall meetings, participates in community cleanup events, and enjoys attending local festivals. Aaron is passionate about fostering a more vibrant version of San Jose and wants to inspire his peers to participate in civic discussions and activities. One of his goals is to create visual simulations of potential urban improvements to share his visions on social media and engage his community. However, Aaron faces a variety of challenges, such as navigation in areas with unsafe pedestrian paths and limited public transportation, the slow pace of change in addressing reported issues, and the underrepresentation of voices of younger residents in urban planning decisions. Since Aaron is proficient with digital tools like social media platforms and mobile apps, he uses the SJ Urban Impact app to report issues that he encounters during daily activities around the city, such as commuting to his part-time job or hanging out with friends. Aaron also utilizes SJ Urban Impact to create visualizations of his urban improvement visions and share them online to spark discussions and gather support for developmental initiatives. He also follows updates on urban development plans and local news by using the news summarization feature on the app.

3.2 Features

The application is designed to address three barriers of civic engagement among young people: difficulty in reporting community issues, difficulty in staying informed about city news, and lack of opportunity to envision the future of the community. To address these issues, SJ Urban impact is designed to analyze and report urban challenges captured in images, provide summaries and key points of lengthy news articles to users, and generate images of users' ideal cities.

Smart Reporting. To support young citizens to report urban issues to local authorities, the ***"Smart Reporting"*** feature (Fig. 1) allows users to simply take a photo and receive a draft report to edit and submit. With this feature, users can submit an image of any issues they encounter in their city and the system will give a detailed analysis of what the issue in the image could be and how it contributes as an urban challenge then gives suggestions on how to address and solve this issue. Once the analysis has been completed and is shown to the user, the app displays a button, allowing users to report this issue to local authorities and alert them of the problem. This feature is powered by the computer vision model in GPT-4, which analyzes the content of the photo. With the text generation model, the system then drafts a brief report of the issue and provides suggestions on how to address the issue. Though we have not implemented the actual reporting functionality, this feature also offers a way for users to theoretically report the issue to local authorities via a submit button.

News Digest. To accommodate the time constraints and information overload that young people face, which prevents them from reading all local news articles, the ***"News Digest"*** feature (Fig. 2) leverages text generation to summarize the key messages in the latest news regarding urban development or community news. From two drop-down menus, users can first select an article or source to read about and then they can choose whether they want to receive a summary, a list of key points, or a full transcription of the writing. By staying updated with ongoing projects and issues in their city, users can be

more informed and stay involved in their communities, by speaking on or taking action on pressing issues they are passionate about.

Feature Envisioning. Finally, to address the lack of active envisioning of the future communities, the ***"Future Envisioning"*** feature (Fig. 3) generates images of urban improvement ideas. It provides two methods for users to envision their ideal city. Drawing on the capabilities of OpenAI's DALLE-3 model, users can select options from a drop-down menu or input a text prompt to visualize urban improvements and other features they would like to see in their city. This gives users an opportunity to bring their creative ideas to life and generate a tangible piece that they can share to others through social media or town hall meetings.

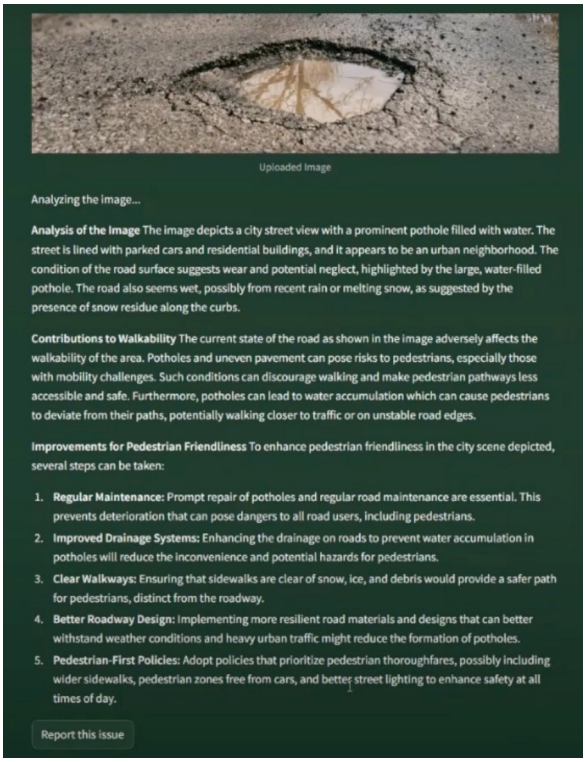


Fig. 1. Smart Reporting

4 Pilot Study and Preliminary Findings

We conducted 45-min interviews and collected pre- and post-study surveys from six student participants in the San Jose area, each with various fields of study (e.g. management information systems, computer science, graphic design, social work). The participants

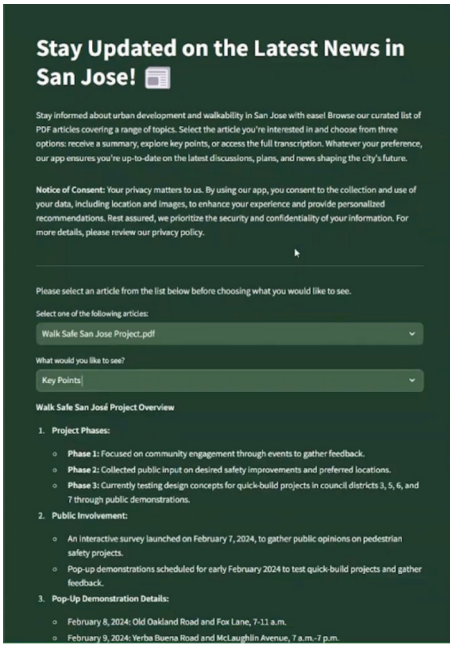


Fig. 2. News Digest

were all full-time juniors, with some working part-time outside of school. In this ongoing research, we identified the themes of the interview data and focused on discussing findings of how generative AI may be useful in encouraging young adults to participate in their communities.

4.1 Simplifying How Young Adults Can Stay Up-to-Date

The majority of participants found great use out of the summarization and key point extraction capabilities of generative AI as they considered 'News Digest' to be the most useful among the three features. Rather than having to search through multiple news sources for location-specific news, having a central hub with the latest information makes it easier for busy individuals to know what is happening in their communities. One student (P2) also appreciated the options available for receiving the news, since she "like[s] reading it in this [key point] format better than one big [paragraph]." Whether users want a summary in paragraph form, key points in an organized list format, or the full transcription of the article for further information, they can choose a format that best suits their needs.

As brought up by one participant (P4), with social media making it easy to get distracted, essentially shortening most people's attention spans, this feature makes it useful for those who may struggle retaining information during long news-reading sessions. Using generative AI to help users digest news offers a balance between keeping it short enough to avoid information overload while also providing adequate information with a sufficient amount of detail. One participant (P3) even suggested being able to input

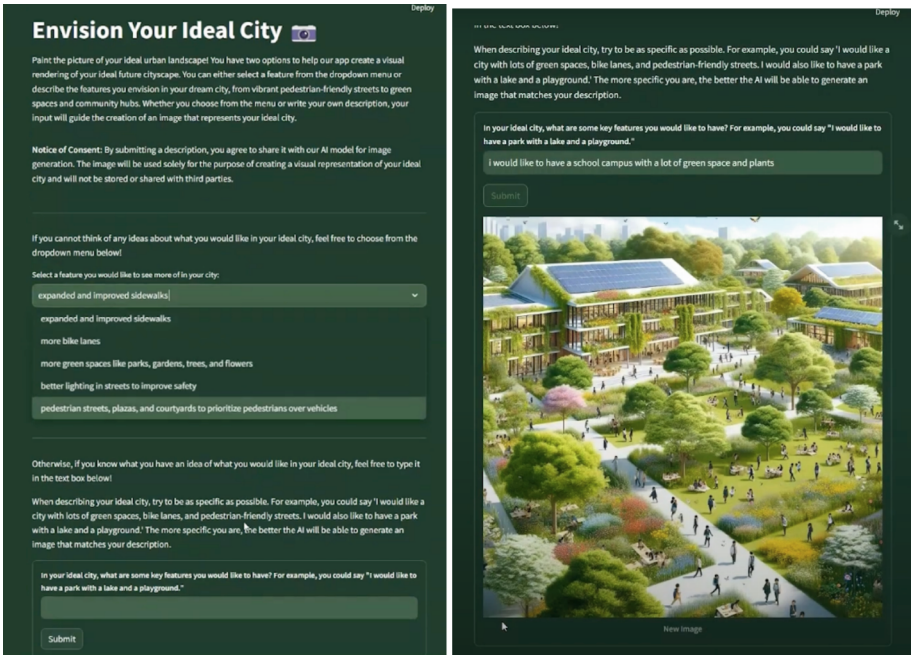


Fig. 3. Future Envisioning

your own news piece instead of choosing from a confined list of articles. She explained, "You have the feature where you could select one of the articles [but] maybe if you could insert your own article and [the AI] can read it somehow, that would be a cool feature." This would allow for users to leverage Gen AI to learn more about news that piques their interest and is more relevant to their unique concerns.

4.2 Streamlining the Reporting Process for Local Issues

Other than using AI to simplify staying up-to-date with the news, many participants also found the Smart Reporting feature to be a quick and easy way to get involved without having to spend too much time and effort, especially considering how busy young adults are as they tend to work and school among other obligations. They commented on the ease and simplicity of being able to take a picture on their mobile devices, upload it even if it is not crystal clear, hit 'Submit', and go on about their day without having to talk to someone about the issue. Furthermore, if more users report the same issue to authorities, they feel that this will increase the urgency and motivation to resolve the issue at hand, which gives them a chance to feel more heard when speaking up or taking action regarding community issues.

However, some participants had ideas on how to improve this feature and make it more interactive. While some issues can be captured in an image and uploaded, some problems are harder or impossible to visualize and can only be discussed, either in verbal or written form. One participant (P6) took it a step further to suggest the ability to chat

with the app. She explained that the generative AI capabilities could be used as "more of a chat […] where the user can ask more questions about it instead of just getting the final output," in order to offer additional information, ask questions for clarification, and finally generate a report that is best suited to their liking before submitting it.

4.3 Offering Feasible Suggestions to Motivated Individuals

While the Smart Reporting feature leverages image analysis and text generation to offer suggestions on how to improve the issues depicted in the uploaded images, participants found the usefulness of the suggestions to be limited. The students were impressed by the ability of the generative AI to analyze an image of an issue, even if it is low-quality, discuss the challenges imposed, explain the severity of the issue, and finally offer ideas on how to improve the issue at hand. On the other hand, since the Smart Reporting feature analyzes the image with the end-goal being a submission of a report to local authorities, the improvement suggestions are tailored towards those who have the ability and resources to make change in the community. Participants felt that the suggestions were not something they could implement themselves. For example, one participant (P4) brought up a good point about using this feature for damaged roads. After submitting an image of a pothole, she received suggestions to regularly maintain roads, use better road-building materials, and improve drainage systems, all of which are not feasible courses of action for a young adult in the general public. She asked "what […] am I supposed to do? I can't fix a pothole."

Regarding this limitation, another participant (P5) suggested having two tiers of suggestions. He explained that once an image has been it uploaded, "first, it should show what the user can do [and call those] 'individual improvements' then the second option would be 'how can the city help[?]" This way, the generative AI can create ideas for what the user can actually do in addition to giving city-level suggestions, which show how local authorities or agencies can address the issue. This gives more flexibility to the user in terms of how they want to use the ideas for improvement, whether they want to do something themselves or gather ideas to present to authorities. This would also provide plenty of feasible options for motivated individuals to take action for their community.

4.4 Providing Customized Volunteer Opportunities

Participants also had other ideas to leverage the capabilities of Generative AI to provide customized volunteer opportunities based on the user's location, interests, schedules, and concerns. Many of the users expressed interest in wanting to participate in their communities through volunteer work, although there are a few obstacles that stand in the way. For many young adults who have to juggle school, work, and personal obligations, this leaves little ample time for volunteering, figuring out who to contact for opportunities, and staying actively engaged in community activities. The barriers increase for those who are not able to drive or find reliable transportation as well as those who relocate often, making it more difficult to find suitable opportunities.

To address these barriers, one participant (P4) proposed, "Tying into the location thing, [the app] knows where […] you are. Maybe it could be like, 'Since you're in this location, this event is happening.' […] What if you have something that curates what

you want?" She suggested that with GenAI in combination with other information like the user's current GPS location on their mobile device, their areas of interests, their concerns, and their calendar schedules, the app can provide a list of curated volunteer events for each user. For each volunteer opportunity, the GenAI can provide general information about the event, a description of the event, the event holder's contact information, and directions/ways to get there. This will eliminate most of the previously mentioned barriers that young adults face when finding volunteer opportunities.

5 Limitations and Future Work

Since SJ Urban Impact is a prototype used for a preliminary study, the app has a few technical limitations and user experience shortcomings. As brought to light by participants in the user study, the app could further incorporate real-time location data into the text and image generation features, so SJ Urban Impact is only limited for users in San Jose, California. Furthermore, for the news feature, the app has a limited number of news articles in its archive because it is not directly linked to the web or any news outlets to get the most current events. As an app that aims to keep young adults involved in their cities, SJ Urban Impact lacks the option to inform users of any ongoing volunteering events or community discussions that are catered to the users' unique interests. Outside the scope of the app, a limitation for image generation is the limited ability to produce realistic visualizations of urban improvement ideas that users have as many of the images created are unfeasible or impractical to implement. As we continue to conduct more user interviews, more areas of improvement will come to our attention for future versions of SJ Urban Impact.

6 Conclusion

This project aims to understand how generative AI can support young adults' civic engagement by prototyping a generative-AI-powered application called SJ Urban Impact. Our pilot study shows that generative AI technologies, such as text generation and image generation, hold the potential to help young citizens stay involved and actively participate in community issues by making this experience more convenient and playful while demanding less effort and time from users. This could potentially reduce barriers and enhance motivation for young people to actively participate in civic issues in their communities. We intend for SJ Urban Impact to be a handy, easily accessible tool for young adults to learn about current events in their community and empower them to speak up about community issues. In the future, we will continue to analyze the interview data and implement user suggestions to address limitations and improve the app, whether it be the app's UI, or functionality. Furthermore, we plan to conduct user studies to evaluate the effectiveness of the app with the new improvements.

Acknowledgments. This project is sponsored by the National Science Foundation under Grant #2142783.

References

Atwell, M., Bridgeland, J., Levine, P.: Civic deserts: America's civic health challenge. In: National Conference on Citizenship. National Conference on Citizenship (2017). https://ncoc.org/wp-content/uploads/2017/10/2017CHIUpdate-FINAL-small.pdf

Bruehlmen-Senecal, E., et al.: Teen and young adult perspectives on generative AI. HopeLab (2024). https://assets.hopelab.org/wp-content/uploads/2024/05/Teen-and-Young-Adult-Perspectives-on-Generative-AI.pdf

Desouza, K.C., Bhagwatwar, A.: Technology-enabled participatory platforms for civic engagement: The case of US cities. Urban Inform., 25--50 (2017)

Hartmann, S., Mainka, A., Stock, W.G.: Opportunities and challenges for civic engagement: A global investigation of innovation competitions. Civic engagement and politics: concepts, methodologies, tools, and applications, pp. 607–623 (2019)

Sharma, A..:11 best generative AI tools and platforms. Turing (2024). https://www.turing.com/resources/generative-ai-tools

Predicting Individual AI Adoption: Personal Technological Innovativeness on Par with AI Attitude

Maike Hering(✉), Sarah Zabel, Nicolas Neef, and Siegmar Otto

University of Hohenheim, Schloss Hohenheim 1, 70599 Stuttgart, Germany
maike.hering@uni-hohenheim.de

Abstract. AI continues to evolve and increasingly permeate our everyday lives. However, the degree to which people adopt AI-based applications varies, raising the question of the underlying drivers. We explored this question in an online study (N = 198 participants) with two data collection waves. First, we assessed potential predictors of AI usage behavior, namely *attitude toward AI* and *personal technological innovativeness*. Second, we assessed the participants' actual AI usage behavior in general and for specific AI applications. A regression analysis revealed that personal technological innovativeness ($b = .18$, $p = .02$ for general AI use, $b = .34$, $p < .001$ for specific AI use), and attitude toward AI ($b = .37$, $p < .001$ for general AI use; $b = 0.25$, $p < .001$ for specific AI use) both predict AI use. We conclude that AI adoption is substantially driven by a general factor – the generic propensity to embrace new technologies (i.e., personal technological innovativeness) – and a technology-specific factor – attitude toward AI. Future research should consider both determinants of AI usage behavior and monitor if, with the continuing adoption of AI in everyday life, personal technological innovativeness loses in importance, while AI attitude wins.

Keywords: AI use · attitude towards AI · Human-Computer-Interaction

1 Introduction

Artificial intelligence (AI) is becoming more ubiquitous every day, but the degree to which people adopt AI-based applications varies substantially. This raises the question of what drives or hinders people to use AI. Our study explores this question by investigating two potential drivers of AI usage: *attitude toward AI* and *personal technological innovativeness* (PTI). Both drivers have shown their potential to predict AI use in the past. However, they are different in nature. While attitudes toward AI are specific to the technology AI, PTI addresses technology adoption more broadly, regardless of the type of technology, focusing more on its novelty [1]. We examine both drivers in an online study with two measurement points and compare their power to predict AI usage in general terms and concerning specific AI applications.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 294–303, 2025.
https://doi.org/10.1007/978-3-031-94171-9_28

2 Drivers of AI Adoption

AI is a technology that evolves quickly, with new AI-based applications constantly being released. This is the basis for our fluid definition of AI and, as we will see below, may affect how attitude toward AI and PTI influence AI adoption.

2.1 What is AI?

When looking at what people **define** as artificially intelligent, we see changes as the technology evolves. The first chatbot in history - ELIZA - sparked a debate about the nature of artificial intelligence because it convincingly mimicked human responses and was considered artificially intelligent when it was created in the 1970s [2]. Nowadays, with the existence of far more sophisticated chatbots such as ChatGPT [3], ELIZA's rather simple, rule-based answers do not appear artificially intelligent anymore. This demonstrates that as AI technology advances, so does our definition of artificial intelligence. In line with the current state of development, we define AI as "any software technology with at least one of the following capabilities: perception—including audio, visual, textual, and tactile (e.g., face recognition), decision-making (e.g., medical diagnosis systems), prediction (e.g., weather forecast), automatic knowledge extraction and pattern recognition from data (e.g., discovery of fake news circles in social media), interactive communication (e.g., social robots or chatbots), and logical reasoning (e.g., theory development from premises)" [4]. Given the dynamic nature of AI, we believe that the factors influencing its usage are also fluid and may change over time. While prior research has examined the two predictors of AI adoption separately [5–7], less is known about how both drivers jointly contribute to AI usage.

2.2 Personal Technological Innovativeness

In short, personal innovativeness can be described as an individual's tendency to "try out new things" [1]. It is considered a trait, meaning that it remains rather stable over the course of life [7]. Applied to the domain of technology, it reflects people's tendency to adopt, experiment with, and embrace emerging technologies [1]. Consequently, studies found that people with high PTI are prone to accept and use AI-based applications, including clinical decision support systems [6], chatbots [5], and driverless buses [7]. Since PTI refers to an individual's tendency to adopt **novel** technologies, it should be related to AI use due to a greater propensity to explore and experiment these emerging technologies.

2.3 Attitude Toward AI

While PTI is a general tendency to embrace new technologies, attitudes toward AI are specific to AI technology. Attitudes are evaluative responses to a particular object (e.g., AI) in affective, behavioral, and cognitive terms [8, 9] and they have been regarded as explanans of related behavior [10]. Applied to AI, this means that AI attitude should be related to AI use. For example, Park et al. [11] report that attitude toward AI drives AI use in a work setting.

Conceptually, attitude toward AI should not be influenced by the novelty of the technology, unlike PTI. Rather, it reflects individuals' evaluation of what they consider to be "AI". We thus expect that both constructs contribute to AI use through different pathways: PTI as a personal trait that facilitates initial engagement with (AI) technologies and attitude towards AI as a determinant of continued use. We, therefore, consider both drivers to gain an understanding of their unique contributions.

2.4 Research Goal and Hypotheses

While previous research has studied the drivers of AI adoption separately, our aim is to test their incremental contributions to explaining AI usage. We expect that both PTI and attitude toward AI uniquely predict AI use beyond the influence of the other.

3 Methods

3.1 Design and Procedure

An online study with two measurement points was conducted. In the initial data collection (July 30 to August 7, 2024), we assessed participants' attitude toward AI and their PTI. In the second data collection (October 17 to October 29, 2024), we captured their AI usage. Data were collected via SoSciSurvey (https://www.soscisurvey.de/) and answered by members of the prolific panel (https://prolific.com). Participants gave informed consent to participate. The data collection for this study was combined with the data collection for a different research project.

The median completion time for the first measurement point was 12 min and 35 s. Participants received £1.82 for their participation. The median completion time for the second survey was 1 min, and participants received £0.16 for their responses.

3.2 Sample Characteristics

In total, $N = 194$ participants took part in both assessments. From this sample, we removed $n = 3$ participants for implausibly fast responses [12], and n = 1 for duplicate participation, retaining the first response. Our sample is representative of the U.S. population (with respect to gender, age, political affiliation, and ethnicity). Please find an overview of the socio-demographic characteristics of the final sample ($N = 190$) in Table 1.

3.3 Measures

If not otherwise stated, all items were answered on 5-point Likert scales with the answer options 1 = strongly disagree, 2 = rather disagree, 3 = neutral, 4 = rather agree, and 5 = strongly agree".

Table 1. Socio-demographic characteristics of the sample.

Variable	Range	Mean	*SD*
Age	18–83	50.03	15.44
		n	%
Gender	Female	90	47.34
	Male	99	52.07
	Diverse	1	0.53
		n	%
Education	Did not graduate from high school	4	2.11
	High school graduate (or equivalent)	21	11.05
	Some college or 2-year degree	66	34.74
	Bachelor's in Computer Science (CS)	8	4.21
	Bachelor's in a different area	59	31.05
	Master's in CS or a similar area	3	1.58
	Master's in a different area	23	12.11
	PhD or higher academic title in CS or a similar area	2	1.05
	PhD or higher academic title in a different area	4	2.11

Attitude Toward AI. Attitude towards AI was assessed with the General Artificial Intelligence Attitude Scale (GAIA-15; [13]). The scale captures people's evaluative response to AI and covers affective, behavioral, and cognitive aspects. Exemplary items are: "AI is exciting" (affective), "I would ride in an autonomously driving car" (behavioral), and "I believe that AI will improve my life" (cognitive). For the analyses, we reversed all negatively formulated items. The GAIA-15 is an item response theory (IRT) scale and to determine participants' AI attitude scores, we conducted a Rasch calibration. This step requires a binary data format, which is why we dichotomized the GAIA-15 item responses before starting the Rasch calibration (1 to 3 were recoded as 0, and the answers 4 and 5 were recoded as 1. This common procedure has been used frequently [13–15].

The Rasch calibration of the GAIA-15 was successful, and the model fit was acceptable with a *Root Mean Square Error of Approximation* of .07, a *Tucker-Lewis Index* of .94, and a *Comparative Fit Index* of .94. The calculated model further showed good empirical reliability with $R = .85$, indicating a high level of precision in estimating the attitude parameters.

Innovativeness. We assessed PTI with the "Innovativeness" subscale from the *Technology-Readiness-Index* [1]. The subscale consists of 4 items and an exemplary item is: "In general, you are among the first in your circle of friends to acquire new technology when it appears".

The scale showed good reliability with a *Cronbach's Alpha* of $\alpha = .85$.

AI Adoption. We captured AI adoption in general terms and with respect to specific types of AI use. To assess General AI Usage, we applied the single item "How often did you use AI applications in the last two weeks?" with the response format *1 = not at all, 2 = rarely (e.g., once or twice in two weeks), 3 = occasionally (e.g., once or twice a week), 4 = often (e.g., every other day), 5 = very often (e.g., daily or more).*

To capture Specific AI Usage, we formulated 7 use cases and assessed how often participants used the concrete AI applications in the last two weeks. Answers were given on the same response scale as for General AI Usage. For the analyses, a mean score was calculated across the 7 items. The use cases were: Voice Assistants (e.g., Siri [16]), AI-based apps to support decisions (e.g., for medical advice), AI-based trading systems to predict stock market prices (e.g., robo advisors), AI copilots for the browser (e.g., Microsoft Copilot [17]), conversational agents (e.g., Claude [18]; or Replika [19]), use of generative AI (e.g., ChatGPT [3]) as an intellectual sparring partner, and AI image generation (e.g., Midjourney [20]). The selection of use cases was based on the definition of AI as a technology that can execute tasks that involve perception (use case 1: voice assistants), decision-making (use case 2: AI-based apps to support decisions), prediction (use case 3: AI-based trading systems), automatic knowledge extraction and pattern recognition from data (use case 4: AI copilots; use case 7: AI image generation), interactive communication (use case 5: conversational agents), or logical reasoning (use case 6: generative AI as intellectual sparring partner) [4].

The items measuring Specific AI Usage showed an acceptable reliability of $\alpha = .77$.

Descriptive Statistics. Means and standard deviations for Innovativeness, General AI Usage, and Specific AI Usage can be found in Table 2.

Table 3 presents descriptive statistics for the Specific AI Usage items, detailing each AI use case individually. It becomes apparent that voice assistants and generative AI are the most frequently used AI applications. AI-based trading systems are the least used AI application, with a mean value close to the lowest response option, suggesting that they are barely used by the general public.

Table 2. Descriptive statistics for Innovativeness, General AI Usage, and Specific AI Usage.

Scale	*N*	*M*	*SD*
Innovativeness	190	3.59	0.96
General AI Usage	190	2.96	1.24
Specific AI Usage	190	1.95	0.77

Note. The response format ranges from 1 to 5, with 1 indicating the lowest level of Innovativeness/AI Usage and 5 indicating the highest level

Correlation Analyses. Participants' GAIA-15 values are IRT person parameters [21]. This means that the units of the GAIA-15 differ from the units of Innovativeness and AI usage. For better comparability, we standardized all variables (GAIA-15, Innovativeness, General AI Usage, and Specific AI Usage) before conducting our correlation and regression analyses.

Table 3. Descriptive statistics for Specific AI Usage at item level.

Use Case	*N*	*M*		*SD*
Voice Assistants	190	2.82		1.63
Decision Support	190	1.69		0.98
Trading Systems	190	1.31		0.76
AI Copilots	188	2.19		1.42
Conversational Agents	185	1.55		0.97
Generative AI	190	2.60		1.39
Image Generation	184	1.52		0.91

Note. The response format ranges from 1 to 5, with 1 indicating the lowest level of AI Usage and 5 indicating the highest level

The results of the correlation analysis are available in Table 4. There are significant correlations between all variables. The highest correlation – a large correlation according to [22] – appears between general and specific AI usage. Descriptively, the GAIA-15 appears to correlate more strongly (medium to large correlation [22]) with General AI usage, while Innovativeness shows a stronger (medium to large correlation [22]) correlation with Specific AI Usage.

Table 4. Correlation Matrix.

	GAIA-15	Innovativeness	General AI Usage	Specific AI Usage
GAIA-15	1	.30 [.16, .43] ***	.41 [.28, .53]***	.34 [.21, .47]***
Innovativeness		1	.33 [.20, .45]***	.44 [.32, .55]***
General AI Usage			1	.68 [.60, .75]***
Specific AI Usage				1

Note. Pearson correlation coefficients (r) are presented with 95% confidence intervals in brackets. All significance tests are two-tailed. All correlations were significant with p-values $p < .001$

4 Results

We conducted our analyses using R Studio in the version 2024.12.0 + 467 [23]. For our analyses, we used the R packages readxl [24], mirt [21], psych [25], tidyverse [26], dplyr [27], and relaimpo [28].

To explore our research questions, we calculated two stepwise regression models with Innovativeness and the GAIA-15 as predictors of General AI Usage (Model 1) and Specific AI Usage (Model 2). For both models, we first used Innovativeness as the sole

predictor and then added the GAIA-15 as a second predictor. The GAIA-15 showed incremental validity in both cases and was maintained in the regression model.

Model 1, predicting General AI Usage based on Innovativeness and the GAIA-15, was statistically significant, $F(2, 176) = 21.79$, $p < .001$, and explained 19.85% of the variance ($R^2 = 0.199$, adjusted $R^2 = .189$). Model 2, predicting Specific AI Usage based on Innovativeness and the GAIA-15, was also statistically significant, $F(2, 176) = 24.28$, $p < .001$, and explained 21.63% of the variance ($R^2 = 0.216$, adjusted $R^2 = .207$). A detailed overview of the regression models' estimates can be found in Table 5.

Table 5. Multiple Regression with the GAIA-15 and Innovativeness Predicting General and Specific AI Usage.

	ß [95% CI]	Std. Error	t value	p
Model 1: General AI Usage				
Intercept	.010 [–.122, .143]	.067	0.154	.878
GAIA-15	.366 [.223, .505]	.071	5.138	< .001
Innovativeness	.177 [.032, .322]	.074	2.400	.017
Model 2: Specific AI Usage				
Intercept	.007 [–.124, .140]	.066	0.113	.910
GAIA-15	.248 [.105, .386]	.071	3.491	< .001
Innovativeness	.345 [.201, .491]	.073	4.696	< .001

We examined the relative importance of the predictor variables and found that Innovativeness accounted for 13% of the variance in Specific AI Usage, while the GAIA-15 explained 8.6%. For General AI Usage, the GAIA-15 explained 14.6% of the variance, while Innovativeness accounted for 5.2%.

5 Discussion

Our analyses confirmed that both attitude toward AI and PTI predict AI adoption. Concretely, a more positive attitude toward AI and greater PTI lead to more frequent AI use. The two constructs are correlated, yet, both independently contribute to predicting AI usage, demonstrating incremental validity in predicting General as well as Specific AI Usage. This suggests that AI usage is driven by a general factor – the generic propensity to embrace new technologies – and a technology-specific factor – attitude toward AI. Our results raise the question of whether this pattern can be generalized and applied to the adoption of other technologies such as robots [29] or self-driving vehicles [30]. If this is confirmed, a two-component theory of technology adoption, consisting of a general and a technology-specific driver, might be appropriate.

5.1 Limitations and Future Research

In this study, AI attitude outperformed PTI in the prediction of general AI usage. When specific AI usage was predicted, however, PTI was a more powerful predictor. Although this effect is descriptively visible when we look at the regression coefficients and the proportion of variance explained by the two predictors, we would like to highlight that the differences were not statistically significant. Nevertheless, we consider them worth interpreting with appropriate caution since they open interesting avenues for future research.

Individuals with high PTI, most likely "early adopters", may have been more familiar with certain AI applications presented in the study (e.g., AI-based trading systems or image generation) and more eager to use them than those with lower PTI. Today, most people will probably have heard of AI and use one or the other AI-based tool (as assessed by our general AI usage item "How often did you use AI applications in the last two weeks?"). In contrast, less common AI use cases may be familiar only to individuals with higher degrees of PTI. Hence, PTI may primarily drive the adoption of less familiar AI technology, while attitude toward AI may play a dominant role in determining the use of more "mainstream" AI applications. Future research should investigate this hypothesis to clarify whether such specific effects of AI attitude and PTI exist.

A similar avenue for future research addresses the developments of the drivers over time. Since PTI reflects individuals' tendency to try out and embrace *emerging* technologies [1] one could argue that its predictive power should decrease over time as AI technology becomes more widely established and less novel. The predictive power of attitude toward AI, however, might not change systematically or even increase somewhat over time since it focusses on the technology itself – which most likely becomes more common – and not on its novelty. Future research should consider both determinants of AI use and assess if, with the continuing adaption of AI in everyday life, PTI becomes less influential compared to AI attitude.

5.2 Implications

AI adoption offers opportunities (e.g., increases in efficiency and productivity or well-being [31, 32]), but also entails significant risks (e.g., high energy consumption or biases in AI systems [4, 33, 34]). To foster a responsible AI use that leverages the benefits and mitigates the risks is therefore key. The finding that PTI and AI attitude drive AI adoption may inform targeted strategies to approach that goal. For example, individuals with a strong positive attitude toward AI and high PTI may be prone to overuse AI and excessively rely on it. Therefore, targeted interventions that raise awareness of AI's potential pitfalls could be especially beneficial for this group. Conversely, people with low AI attitude and limited PTI may benefit from interventions that enhance AI competence, enabling them to start using AI-based tools.

6 Conclusion

In conclusion, our study highlights the dual role of PTI and attitude toward AI in predicting AI adoption. Through an online study with two data collection waves, we found that both factors significantly contribute to AI usage behavior. Our findings align with previous research that underscores the importance of technology-specific attitudes and PTI in predicting technology adoption. Descriptively, we find that PTI plays a more prominent role in predicting the adoption of specific AI applications, while attitude toward AI was a stronger determinant of general AI use. This opens an avenue for future research, which should explore whether this is a systematic pattern. Additionally, given the novelty-driven nature of PTI, another promising research direction is to examine whether this factor in AI adoption steps back behind AI attitude as AI technology becomes more established.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this Article.

References

1. Lam, S.Y., Chiang, J., Parasuraman, A.: The effects of the dimensions of technology readiness on technology acceptance: an empirical analysis. J. Interact. Mark. **22**, 19–39 (2008). https://doi.org/10.1002/dir.20119
2. Natale, S.: If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. New Media Soc. **21**, 712–728 (2019). https://doi.org/10.1177/1461444818804980
3. OpenAI: ChatGPT (2025)
4. Vinuesa, R., et al.: The role of artificial intelligence in achieving the Sustainable Development Goals. Nat. Commun. **11**, 233 (2020). https://doi.org/10.1038/s41467-019-14108-y
5. Tian, W., Ge, J., Zhao, Y., Zheng, X.: AI Chatbots in Chinese higher education: adoption, perception, and influence among graduate students-an integrated analysis utilizing UTAUT and ECM models. Front. Psychol. **15**, 1268549 (2024). https://doi.org/10.3389/fpsyg.2024.1268549
6. Yu, Z., et al.: The willingness of doctors to adopt artificial intelligence-driven clinical decision support systems at different hospitals in china: fuzzy set qualitative comparative analysis of survey data. J. Med. Internet Res. **27**, e62768 (2025). https://doi.org/10.2196/62768
7. Chen, J., Li, R., Gan, M., Fu, Z., Yuan, F.: Public acceptance of driverless buses in China: an empirical analysis based on an extended UTAUT model. Discret. Dyn. Nat. Soc. **2020**, 1–13 (2020). https://doi.org/10.1155/2020/4318182
8. Smith, A.N.: The importance of attitude in foreign language learning. Mod. Lang. J. **55**, 82 (1971). https://doi.org/10.2307/321854
9. Rokeach, M.: A Theory of Organization and Change Within Value-Attitude Systems1. Journal of Social Issues, vol. (1968). https://doi.org/10.1111/j.1540-4560.1968.tb01466.x
10. Garms-Homolová, V.: Einstellung: Begriff, Konzepte und praktische Bedeutung. In: Garms-Homolová, V. (ed.) Sozialpsychologie der Einstellungen und Urteilsbildung. Psychologie für Studium und Beruf, pp. 1–14. Springer, Heidelberg (2020). https://doi.org/10.1007/978-3-662-62434-0_1#DOI
11. Park, J., Woo, S.E., Kim, J.: Attitudes towards artificial intelligence at work: Scale development and validation. J. Occupat. Organ. Psyc. **97**, 920–951 (2024). https://doi.org/10.1111/joop.12502

12. Leiner, D.J.: Too Fast, too Straight, too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys. 229–248 Pages / Survey Research Methods, vol 13 No. 3 (2019). SRM, vol. 13, pp. 229–248 (2019). https://doi.org/10.18148/srm/2019.v13i3.7403
13. Hering, M., Neef, N.E., Zabel, S., Otto, S.: General Artificial Intelligence Attitude (GAIA) Item Pool – How to Measure Attitudes Toward an Ever-Changing Object. Center for Open Science (2024). https://doi.org/10.31235/osf.io/3c8sr
14. Neef, N.E., Zabel, S., Otto, S.: Climate change mitigation: a question of humanitarian or environmental motivation? J. Environ. Psychol. **100**, 102483 (2024). https://doi.org/10.1016/j.jenvp.2024.102483
15. Kaiser, F.G., Byrka, K., Hartig, T.: Reviving CAMPBELL'S PARADIGM FOR ATTITUDE RESEARCh. Pers. Soc. Psychol. Rev. **14**, 351–367 (2010). https://doi.org/10.1177/1088868310366452
16. Apple Inc.: Siri, vol. https://www.apple.com/de/siri/ (2025)
17. Microsoft: Microsoft Copilot (2025)
18. Anthropic: Claude (2025)
19. Replika: Replika (2025)
20. MidJourney: Midjourney (2025)
21. Chalmers, R.P.: mirt: a multidimensional item response theory package for the R environment. J. Stat. Soft. **48** (2012). https://doi.org/10.18637/jss.v048.i06
22. Cohen, J.: Statistical power analysis for the behavioral sciences. L. Erlbaum Associates, Hillsdale N.J. (1988)
23. R Core Team: RStudio: Integrated Development for R. Vienna, Austria (2020)
24. Wickham, H., Bryan, J.: readxl: Read Excel Files (2023)
25. Revelle, W.: psych: Procedures for Personality and Psychological Research. Evanston, IL (2021)
26. Wickham, H., et al.: Welcome to the Tidyverse. JOSS **4**, 1686 (2019). https://doi.org/10.21105/joss.01686
27. Wickham, H., François, R., Henry, L., Müller, K.: dplyr: A Grammar of Data Manipulation. R package version 1.0.7 (2021). https://CRAN.R-project.org/package=dplyr
28. Grömping, U.: Relative importance for linear regression in R: The Package relaimpo. J. Stat. Soft., vol. 17 (2006). https://doi.org/10.18637/jss.v017.i01
29. Neef, N.E., Zabel, S., Lauckner, M., Otto, S.: What is appropriate? on the assessment of human-robot proxemics for casual encounters in closed environments. Int. J. Soc. Robot. (2023). https://doi.org/10.1007/s12369-023-01004-1
30. Neef, N.E., Kastner, K., Schmidt, M., Schmidt, S.: On Optimizing driving patterns of autonomous cargo bikes as a function of distance and speed—a psychological study. IEEE Open J. Intell. Transp. Syst. **3**, 592–601 (2022). https://doi.org/10.1109/OJITS.2022.3198120
31. Mills, S., Spencer, D.A.: Efficient Inefficiency: Organisational challenges of realising economic gains from AI. J. Bus. Res. **189**, 115128 (2025). https://doi.org/10.1016/j.jbusres.2024.115128
32. Maples, B., Cerit, M., Vishwanath, A., Pea, R.: Loneliness and suicide mitigation for students using GPT3-enabled chatbots. Npj Mental Health Res. **3**, 4 (2024). https://doi.org/10.1038/s44184-023-00047-6
33. Zabel, S., Otto, S.: SustAInable: how values in the form of individual motivation shape algorithms' outcomes. In: An Example Promoting Ecological and Social Sustainability. Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA (2024). https://doi.org/10.1145/3613904.3642404
34. Zabel, S., Otto, S.: Bias in, bias out – The similarity-attraction effect between chatbot designers and users. In: International Conference on Human-Computer Interaction. LNCS, vol. , pp. 184–197 (2021). https://doi.org/10.1007/978-3-030-78468-3_13

Decentralized Systems and AI: Exploring the Intersection of Web3, Blockchain, and Predictive Machine Learning

Myriam Hernández-Alvarez[1(✉)] and Edgar Torres-Hernández[2]

[1] Escuela Politécnica Nacional, Quito, Ecuador
myriam.hernandez@epn.edu.ec
[2] Invex Negocios, Quito, Ecuador

Abstract. This paper explores the integration of decentralized systems, Web3, and predictive machine learning in a case study for the development of a job recruitment platform. By leveraging the Internet Computer Protocol (ICP), the platform ensures secure, transparent, and autonomous hiring processes without reliance on centralized intermediaries. The architecture incorporates blockchain-based identity verification, smart contracts for immutable job transactions, and AI-driven candidate matching. The study evaluates the system's performance through both qualitative and quantitative methodologies, measuring latency, scalability, accuracy, and user trust. While the decentralized nature enhances data security and user autonomy, challenges such as system latency and user onboarding remain. The findings suggest that decentralized AI-driven recruitment systems offer significant advantages in privacy protection and security but require further optimization for mainstream adoption. Future research will explore federated learning, enhanced scalability, and broader applications in freelancing and gig economies.

Keywords: Decentralized Systems · Web3 · Blockchain · Predictive Machine Learning · Job Recruitment

1 Introduction

The rise of Web3 technologies and blockchain has enabled the development of decentralized platforms that enhance security, privacy, and autonomy. This paper explores the integration of decentralized architectures and predictive machine learning in a job recruitment system. By leveraging Internet Computer Protocol (ICP) [1] and smart contracts, this project aims to implement job matchmaking in the blockchain while ensuring user data ownership and privacy.

To demonstrate the viability of decentralized AI applications, we present a proof-of-concept implementation of a decentralized employment board. This system enhances security by distributing data across ICP canisters, eliminates reliance on centralized entities, and utilizes machine learning models to optimize job matching.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 304–311, 2025.
https://doi.org/10.1007/978-3-031-94171-9_29

1.1 The Internet Computer Protocol (ICP)

The Internet Computer Protocol (ICP) represents a groundbreaking advancement in decentralized application development, offering a scalable and efficient blockchain-based infrastructure beyond traditional smart contracts. Developed by the DFINITY Foundation, ICP eliminates reliance on centralized cloud services by enabling direct deployment and execution of decentralized applications (dApps) [2]. With its ability to operate at web speed, ICP facilitates transactions and computations within milliseconds, significantly outperforming legacy blockchains like Ethereum. This rapid processing is achieved through chain key cryptography, which enhances validation efficiency, reduces network congestion, and lowers operational costs. Additionally, ICP's infinite scalability allows developers to create complex applications without facing computational bottlenecks or excessive fees [3].

A fundamental feature of ICP is its autonomous canister architecture, which acts as a decentralized alternative to traditional cloud-based storage and computing. These self-contained computational units securely store data, execute smart contract logic, and interact efficiently with other canisters, ensuring enhanced security, reduced latency, and cost-efficient execution. By maintaining a trustless and censorship-resistant environment, ICP removes dependence on centralized entities, making it ideal for industries prioritizing data privacy and security, such as decentralized finance (DeFi), supply chain management, and decentralized job recruitment platforms.

Recent advancements in ICP have further integrated artificial intelligence (AI) into its decentralized framework. The deployment of DeepSeek's AI models on the ICP platform enhances dApps by providing advanced AI-driven functionalities within a decentralized environment [4]. Collaborations with projects like ICPanda DAO have enabled the incorporation of AI agents, expanding the capabilities of dApps. Moreover, tools like CaffeineAI facilitate AI-driven software development within ICP, allowing AI to autonomously build and deploy websites or applications directly on-chain via smart contracts. With canisters capable of handling substantial data, ICP enables efficient and scalable AI-driven software solutions.

By combining speed, scalability, and security, ICP serves as an optimal foundation for next-generation Web3 applications, particularly in decentralized recruitment systems that require secure identity verification, tamper-proof data storage, and AI-powered candidate matching. With its continued evolution and AI integration, ICP will drive innovation across multiple industries, reinforcing its position as a leading decentralized computing protocol.

2 Introduction

2.1 The Internet Computer Protocol (ICP)

The platform consists of three components. The front end serves as the user interaction layer, providing a decentralized Web3 Dapp that allows users to interact with the system. The profiles of candidates and recruiters are securely stored in individual canisters for privacy. Employers post job listings while candidates apply securely, and an AI-powered recommendation engine uses AI algorithms to match candidates with jobs based on qualifications and preferences.

The backend infrastructure handles decentralized computation and storage where each user has a separated ICP canister to store resumes, job applications, and credentials. Smart contracts enable transactions such as job postings, applications, and updates. The system retrieves and processes data from the canisters, maintaining decentralization.

Predictive machine learning integration enhances job matching through a candidate-matching algorithm that analyzes job descriptions, resumes, and preferences to rank job-candidate compatibility. Distributed model training allows AI models to train on decentralized datasets without the need for centralized storage, and privacy-preserving AI ensures that machine learning inferences are performed within canisters to minimize data exposure.

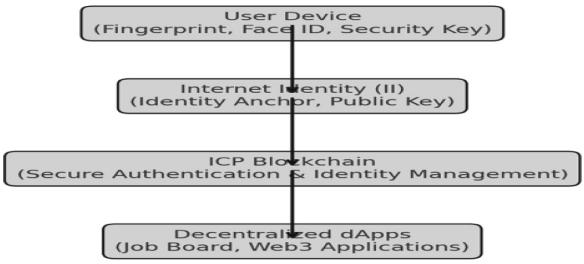


Fig. 1. Internet Identity Process used in ICP.

2.2 System Workflow

User registration and identity verification begin with decentralized identity authentication, where users store their credentials in personal canisters. Internet Identity (II) in Internet Computer Protocol (ICP) provides a decentralized, privacy-preserving, and secure authentication mechanism that removes the need for traditional usernames and passwords, as shown in Fig. 1. By using cryptographic keys tied to their devices; users can securely log in to decentralized applications (dApps).

Instead of traditional centralized authentication databases, ICP's Internet Identity leverages WebAuthn and cryptographic signatures to verify user identities. Users can be authenticated with biometric sensors from a movil device such as fingerprint or Face ID or use hardware security keys like YubiKey and Ledger to eliminate the need for passwords to avoid phishing and credential theft risks. Additionally, each application generates a unique pseudonymous identity for the user, preventing tracking across services. Notably, authentication is fully decentralized, with no reliance on a central entity.

The authentication process consists of three steps, presented in Fig. 2. In user enrollment; the user registers a device such as a laptop, phone, or security key to generate a cryptographic key pair, with the public key stored in a secure identity anchor on the ICP blockchain. During login, the user selects Internet Identity as their authentication method, and their device signs a challenge using the stored key pair. The identity is verified locally without exposing credentials. Access control and permissions ensure that

each dApp assigns a unique, unlinkable identity per user, preventing correlation between identities across different applications, and thereby preserving privacy.

The benefits of Internet Identity for the decentralized job board include improved security, as it protects against brute-force attacks and phishing scams. This mechanism also enhances data privacy because the users can manage their identities without revealing personal details. The user experience remains simple through biometric login, making authentication easy for non-technical users, while cross-device compatibility enables secure authentication across multiple devices.

Internet Identity in ICP is a next-generation authentication method that ensures security, privacy, and decentralization [5]. In our case, it allows the users full control over their credentials while maintaining seamless and secure platform access.

Employers submit job postings on-chain while candidates update their profiles, ensuring data remains secure. The proposed approach ranks matches based on job requirements and candidate profiles. The user applications for jobs and the hiring requests are managed via smart contracts. The system also tracks performance and collects user feedback to improve AI matching through qualitative and quantitative evaluations.

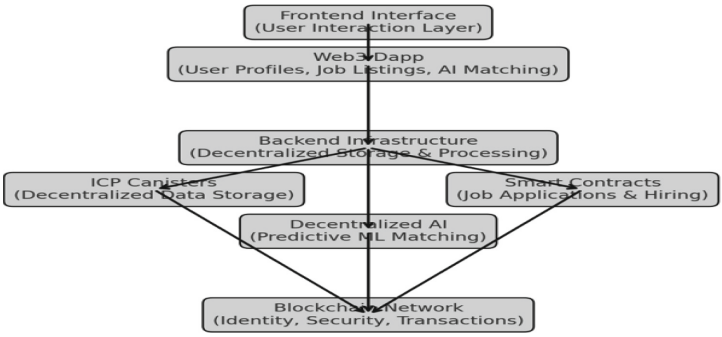


Fig. 2. System Workflow.

2.3 Security and Privacy Considerations

User-controlled data ensures that candidates retain full ownership over their CVs and application history. Decentralized authentication enhances security by reducing the risk of impersonation, as identity verification occurs directly on the blockchain. By distributing storage across multiple canisters, the system prevents large-scale data breaches, reinforcing security and privacy at every level of the recruitment process.

3 Methodology

We used a hybrid methodology for evaluation, combining quantitative performance metrics and qualitative user feedback. We conducted experiments in a controlled environment to measure key factors such as latency, scalability, and accuracy. Latency was

assessed by evaluating the system's response time for decentralized job searches and AI-driven recommendations. Scalability was analyzed by observing performance under an increasing number of users and job applications. Accuracy was determined by examining the effectiveness of the AI model in predicting relevant job matches.

In addition to these quantitative assessments, qualitative evaluations were conducted through focus groups and surveys with students and recruiters analyzing usability measured by the ease of navigating the platform, trust by gauging user confidence in the security and accuracy of decentralized AI-based recruitment, and also we evaluated privacy awareness by determining the level of understanding regarding decentralized data ownership among users.

4 Implementation and Results

4.1 Technical Implementation

The decentralized recruitment platform is called ICP Jobs, and it was implemented using Motoko, the programming language for ICP smart contracts [6]. The machine learning model was deployed in a decentralized execution environment. The implementation included three aspects: storage in individual canisters for the information of each user, on-chain transactions handled through smart contracts, and AI-powered recommendations using machine learning algorithms to encounter the best matches between hiring requests and job applications.

ICP Jobs leverages ICP to manage student CVs and job requirements. By using the ICP's canister model, our application securely stores and retrieves data while ensuring that only authorized users can perform certain actions (for example, deleting all records). The project is implemented using Motoko for the backend and standard web technologies for the front end. The code structure is shown in Fig. 3.

The key features include actions related to CVs, such as adding, viewing, searching by skills, managing job requirements, and deleting records. The backend logic is developed in Motoko. ICP Jobs was deployed as a proof-of-concept to demonstrate how ICP can be used to build decentralized applications with secure, permissioned interactions. By integrating a robust backend in Motoko with a responsive frontend, our project highlights the potential for decentralized management of digital credentials and job opportunities.

4.2 Quantitative Findings

The decentralized system exhibits an average latency of 150 ms, and the centralized baseline has a faster response of 50 ms for similar processing. This increased latency is due to the overhead associated with decentralized data retrieval and processing in several canisters requiring additional network communication and consensus mechanisms related to the blockchain environment. While this additional processing results in slower response times compared to centralized systems, the trade-off is a significant enhancement in data security and user privacy.

The performance of the AI-powered recommendation engine in the decentralized system achieves a match accuracy of 87%, which is slightly below the 90% accuracy

```
icp-jobs/
├── dist/                          # Compiled and bundled frontend assets
├── src/
|   ├── icp-jobs-backend/          # Backend canister code in Motoko
|   |   ├── main.mo                # Main Motoko actor for business logic and AI processes
|   |   ├── ai_module.mo           # AI module for candidate-job matching
|   ├── icp-jobs-frontend/         # Frontend application code
|   |   ├── app.js                 # JavaScript to handle UI events and canister interaction
|   |   ├── index.html             # Main HTML file
|   |   ├── styles.css             # Styling for the app
├── declarations/
|   └── icp-jobs-backend.js        # Auto-generated canister declarations for JS interop
├── package.json                   # Project dependencies and scripts
├── webpack.config.js              # Configuration for bundling frontend assets
├── dfx.json                       # DFINITY configuration for canister deployment and network settings
└── README.md                      # Project documentation
```

Fig. 3. Code Structure.

of the centralized baseline, probably because in a decentralized environment, data is partitioned across multiple canisters, potentially leading to less diverse training data for the machine learning algorithms. However, the 87% accuracy is a good result that indicates that the system remains effective with the added enhanced privacy.

The most important characteristic of the decentralized approach is the lower risk of security breaches [7], in contrast to the centralized systems that can be attacked in their data repository, affecting all the information. In contrast, in the proposed architecture, the data is stored in separate canisters, which are protected with cryptography. This distribution of the data reduces the attack surface, and even if one canister is compromised, the breach would be isolated, preventing large-scale data theft. Conversely, centralized systems store data in unified databases, making them attractive targets for cyberattacks that can result in extensive data breaches.

In conclusion, as shown in Table 1, there is a trade-off between latency and a slight decrease in accuracy versus important improvements in data security and privacy. The trade-offs highlight the inherent challenges and benefits of adopting decentralized architectures for applications like job recruitment platforms.

Table 1. Quantitative Results

Metric	Decentralized System	Centralized Baseline
Latency (avg ms)	150ms	50ms
Match Accuracy (%)	87%	90%
Security Breach Risk	Low	High

4.3 Qualitative Insights

As part of the project, user feedback was collected in a questionnaire. Results showed that the users valued the improved data ownership. However, some interfaces Web3 were

found difficult to understand and navigate, pointing to the need for a more intuitive interface design. The AI recommendations were appreciated for job matching but expressed the need for transparency about how the matches were generated.

Finally, some users expressed a preference for a verification mechanism that would ensure job postings and candidate profiles were legitimate to avoid fraud. As a solution the authors plan a future implementation of decentralized reputation scores and blockchain-verified credentials to increase confidence in the platform.

User trust is critical for widespread adoption. Therefore, future iterations of the software will incorporate these recommendations to obtain enhanced explainability, verifiable identity mechanisms, and interactive user guidance to foster greater confidence in the recruitment process.

5 Discussion and Future Work

This decentralized approach in the ICP environment, applied to a recruitment board, demonstrated the benefits of Web3 and AI integration but also presented some challenges to resolve. One of the areas for improvement is latency optimization. To reduce delays in data retrieval and processing, local indexing and off-chain computation could be considered.

Other potential improvements are the design of simplified UI/UX and the delivery of comprehensive educational resources to improve accessibility, particularly for users unfamiliar with decentralized technologies.

By proving that a job board application can be built in a decentralized manner while still providing the ease of use, scalability, and security commonly associated with Software as a Service SaaS models, the present project challenges the conventional paradigm. It shows that decentralized architectures can support practical, real-world applications like job boards that need to be fast, safe, and secure without sacrificing the operational efficiencies and benefits typically reserved for centralized SaaS systems.

In summary, while traditional SaaS models rely on centralized infrastructures, the proposed work leverages the ICP to provide a decentralized alternative that still delivers the benefits of SaaS, making it a proof-of-concept for a new generation of secure, scalable, and user-centric applications.

Looking ahead, there are various directions for future research to refine and expand the capabilities of the system. However, the objective of this work was fulfilled because we proved the viability of the implementation of a system in the ICP environment to test the user experience and find out the challenges that this implementation could face. Furthermore, the decentralized job-matching methodology could be expanded to other contexts that could also benefit from the enhanced privacy and security that offers the blockchain approach.

Acknowledgments. This paper was published thanks to the funds provided by Vicerrectorado de Investigación, Innovación y Vinculación of the Escuela Politécnica Nacional in Quito, Ecuador.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alli, O., Kallach, L.: Artificial Intelligence Enabled Human Resources Recruitment Functionalities: A Scoping Review. Procedia Comput. Sci., 3268–3277 (2024)
2. Li, A., Serena, L., Zichichi, M., Tang, S.K., D'Angelo, G., Ferretti, S.: Modelling of the internet computer protocol architecture: The next generation blockchain. In: International Congress on Blockchain and Applications (2022)
3. Cheng, S.: Applications of Web 3.0. Appl. Web 3.0. Springer Nat. Singapore, pp. 109–145 (2024)
4. Jain, V.: Revolutionizing AI with ICP. https://medium.com/web3-blockchain-tech/revolutionizing-ai-with-icp-9cb574ae5a11
5. Zeydan, E., et al.: Blockchain-based self-sovereign identity: Taking control of identity in federated learning. IEEE Open J. Commun. Soc. (2024)
6. Torres-Hernandez, E., Hernandez-Alvarez, M.: ICP Jobs. https://github.com/PerArduaAdAstra777/ICP-Jobs-New
7. Zyskind, G., Nathan, O., Pentland, A.: Decentralizing Privacy: Using Blockchain to Protect Personal Data. IEEE Secur. Priv. (2015)

Importance of Delegation and Redelegation in Human-Autonomy-Teamed Air Traffic Control

Robert Hunger[1(✉)] and Sabine Theis[2]

[1] German Aerospace Center (DLR), Institute of Flight Guidance, Braunschweig, Germany
[2] German Aerospace Center (DLR), Institute of Software Technology, Cologne, Germany
{robert.hunger,sabine.theis}@dlr.de

Abstract. Across almost all European air navigation service providers there is a current shortage of air traffic controllers (ATCOs), which contributes significantly to delayed or cancelled flights. This reduces the cost efficiency, increases the environmental impact of air traffic and causes inconvenience to airlines and passengers. Looking at the opportunities the widely discussed new generation of artificial intelligence offers to air traffic control one researched solution is the introduction of a digital (artificial intelligence) air traffic controller. It can be expected that the first operational use of a digital ATCO will not be allowed nor able to take over air traffic tasks of ATCOs to their full extent. A researched (interim) solution is a (human autonomy) team build by ATCO(s) and digital ATCO to handle air traffic together in parallel within the same airspace. In this setup, situations can develop that digital ATCO is not (yet) able to handle or where ATCO likes to change allocation as she will stay responsible for safe air traffic control. This requirement for a bi-direction change of authority (i.e. transfer of control) is termed delegation and redelegation in this poster. A first demonstrator of a functional delegation and redelegation process and its implementation was presented to seven ATCOs in a workshop and tested by them using short scenarios to clarify questions about the user interface and user experience. The results showed UEQ's Pragmatic Quality scores of above 1.4 proving the user-friendliness of the delegation process.

Keywords: Digital Controller · Human Autonomy Teaming · Future Air Traffic Control

1 Introduction

The air traffic management system in Europe faces a significant challenge: with recovering and predicted further growing traffic levels [5] the shortfalls in en-route capacity reached the worst level since 20 years in 2023 [6]. The resulting

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 312–322, 2025.
https://doi.org/10.1007/978-3-031-94171-9_30

delays can be assigned to weather, capacity, air traffic control (ATC) disruption and ATC staffing in the order given. However, the latter plays a role in 80% of the cases assigned to capacity and disruption as well. This makes ATC staffing the primary factor for air traffic flow management (ATFM) delays; in 2023 contributing to the total en-route ATFM delay of 18.1 million min [6]. Thus, aircraft flying over Europe were unnecessarily airborne for over 300,000 h causing pollution and inconvenience to passengers. Based on a rough calculation for an A320neo (as one of the most common types of aircraft) with 165 seats and a fuel burn of 0.01758 kg/km/seat [11] and a cruising speed of 450 kn [4] this equals more than 2.28 M kg CO_2. Addressing these issues by researching and implementing higher levels of automation, i.e. introduction of a digital ATCO (D-ATCO) raises followup challenges. Firstly, research has shown that pure automation of ATC while leaving the responsibility to human (supervising) will cause decrease of situational awareness and ATCO skills [3,13] making an effective supervision unrealistic. Secondly, a full automation of ATC is an costly (time, effort, money) undertaking, not to be expected to be implemented in the short term. However, focusing on intermediate steps, this again reinforces the first point. This is where human-autonomy-teaming (HAT) research comes in, enabling an earlier implementation, relieving human (or solving staff shortages and enabling the processing of the demand) and giving possibility to keep ATCO skills and situation awareness high enough to take over [7,17]. In turn questions needs to be answered regarding the parallel work of human and D-ATCO in one airspace. [14] mentions here for example:

- Which flights are allocated to whom and considering what aspects?
- Who is responsible for a conflict between aircraft under ATCO and D-ATCO control?

Whatever the answer is, a possibility for delegation is crucial. While these researches focusing on different aspects of human-autonomy-teaming and addressing them, they always assume a possibility for human to delegate, i.e. give control to or take control back from D-ATCO. They lack a concrete investigate of how this should be conventionalized and implemented. Furthermore, it is limited to the ATCO being actively requesting control back or allocating control to D-ATCO. [14] found that ATCOs will use individual strategies for allocation preventing the effective implementation of general allocation rules. Additionally, this ATCO-driven allocation will cause additional workload; ATCOs in [1] noted and liked to be relived from. However, they still favoured to keep the opportunity to change the allocation. This poster presents related work in Sect. 2 regarding research in digital ATCO in general and human-autonomy-teaming in ATC in particular. This includes previous workshops together with ATCOs as well as presentation of D-ATCO demonstrator system. Section 3 presents the workshop, of which the study and evaluation of the delegation discussed here formed a part. The used metrics and techniques are described. Subsequent to the presentation of results in Sect. 4, they are discussed in Sect. 5.

2 Related Work

A Digital ATCO as an concept for artificial-intelligence-based automation support in ATC was for example drafted by [8] and developed as prototype by [10] in the frame of German Aerospace Centre (DLR)'s LOKI project [2]. Previous Workshops identified ATCO's expectations, most relevant task for human autonomy teaming (HAT) and situations and issues to focus on; including delegation. In conclusion, tasks (or flights) that can be described as simple are delegated to D-ATCO [1]. However, in the dynamic ATC environment a task or flight might become non-simple with changing surrounding conditions.

The Automation Strategy of EURONCOTROL Maastricht Upper Area Control Centre foresees the same workshare between human ATCO and D-ATCO: basic and routine ATCO tasks and traffic allocated to a digital ATCO (named ARGOS), whereas complex scenarios will be handled by ATCO with D-ATCO support. ARGOS is at an early implementation stage in a constrained environment. However, the strategy foresees the possibility for ATCO to take back control [9].

These two example show that not just flight allocation is a crucial part in HAT ATC-environment [15], but especially the possibility to change this initial allocation.

3 Method

This Chapter defines the term delegation as used in this poster before the workshop setup and used metrics are described which were used to verify the delegation (between ATCO and D-ATCO) concept and prototype-implementation for an future HAT ATC environment through ATCO feedback.

3.1 Definitions

Delegation is used e.g. by [15] as a term to describe that human ATCO has currently flights allocated to a D-ATCO. This sets the focus not on the process of delegating, but rather on the status of an flight being under human or D-ATCO control. In contrast delegation (and redelegation) is used as an umbrella term to describe any transfer of control process between ATCO and D-ATCO throughout this poster. However, it is important to define sub-terms dependent on who is acting or reacting and who is sending or receiving an aircraft as the four possible combinations (modalities; capitalised) differ in importance, impact and thus in proposed implementation. The definitions are shown in Table 1.

3.2 Workshop Setup

As followup of the previous two workshops (see Sect. 2) a two-day workshop was conducted in autumn 2024 in Langen, Germany with operational ATCOs. First of all, the workshop was used to introduce the team and participants, the

Table 1. Delegation modalities

Name	Role initiating the transfer	Role receiving the aircraft	Clarifications
Delegation	ATCO	D-ATCO	disabled, if D-ATCO is not able or allowed to handle aircraft
Cancel Delegation	ATCO	ATCO	immediately transferred
Redelegation	D-ATCO	ATCO	time critical; cause information
Request Delegation	D-ATCO	D-ATCO	affirmation leading to Delegation

Fig. 1. Delegation, Redelagation (due to received coordination) and Cancel Delegation Implementation

project and the D-ATCO prototype. Thereafter, participating ATCOs were split in three groups to assess three D-ATCO dedicated functionalities in one hour sessions each. However, this poster is limited to ATCO feedback to concept and implementation of delegation. Within each session three scenarios were shown to each group, which consisted of two or three ATCOs. One of the ATCOs actively handled the system, whereas the other(s) act as bystanders, but they can discuss together. Each scenario focused on one of the modalities Delegation, Redelegation and Request Delegation (see Fig. 1), with Cancel Delegation possible in every one, but explicitly addressed in the Request Delegation scenario. All scenarios were designed for five minutes, however they were able to run longer. The scenarios were shown through a radar display simulation tool imitating DFS's operational iCAS-System. Whereas the radar display was fully functional, neither the D-ATCO nor the allocation process was, as the sessions exclusively focused on the delegation process. Instead aircraft were allocated during scenario design to create best possibilities to test the delegation modality in focus. After each scenario a multiple-choice questionnaire with tailored questions and UEQ-S was answered by the ATCOs, followed by a semi-structured interview. However, depending on time left due to queries and technical issues the discussion length and addressed items of this interview varied between groups.

3.3 Materials

User Experience Questionnaire Short Version (UEQ-S). During the evaluation workshop we assed the user experience using the User Experience Questionnaire - Short Version (UEQ-S) [16], a validated instrument designed for rapid evaluation of interactive systems. The UEQ-S consists of eight bipolar items, each rated on a 7-point Likert scale (1 = very negative, 7 = very

positive). The questionnaire measures two dimensions: Pragmatic Quality (PQ): Evaluates usability, efficiency, and clarity. Hedonic Quality (HQ): Assesses the system's attractiveness and emotional appeal. The rationale for applying the UEQ-S was its efficiency in capturing user experience while maintaining reliability and validity comparable to the complete User Experience Questionnaire (UEQ) [12].

Self-Administered Questionnaires on User Feedback. The study utilized a structured, self-administered questionnaire to assess participants' preferences, perceptions, and usability ratings regarding task Delegation, Redelegation, Cancel Delegation, and Request Delegation when using described prototype. The questionnaire comprised multiple-choice questions, ranking tasks, and Likert-scale ratings. The Delegation questionnaire asked participants first chose their preferred method for delegation function implementation, weighing the risk of unintended delegations. Options included a single click, click and hold, double click or a confirmation pop-up. They also indicated their preference for task delegation instead of aircraft delegation, selecting between individual aircraft assignments, specific tasks for all aircraft in their area of responsibility (AOR), undecided/additional to aircraft delegation, or no task delegation. For observing Delegation completion, participants selected from multiple options for feedback, such as changes in the head symbol, highlighting, additional symbols, hiding label information, or modifications to label color or style. They specified conditions for using Delegation, including routine tasks D-ATCO is qualified for, after resolving conflicts for aircraft that requires no further interaction, following answered pilot requests, in cases of personal workload overload, or never. The Redelegation questionnaire evaluated participants' views on task reassignment in ATC. They indicated if a radio check was needed post Redelegation and ranked various aircraft scenarios by priority (sequence of redelegations), including conflicts and pilot requests. Participants also specified situations where the system (D-ATCO) should redelegate control automatically. The Cancel Delegation questionnaire evaluated participants' needs for information and feedback when canceling a delegated task. Participants indicated if they needed information for cancellation and expressed preferences for observing the cancellation process through various options as already described and used for Delegation questionnaire. Finally, they were asked if a radio check should be performed after canceling the delegation, with options for yes, undecided, or no. The Request Delegation questionnaire evaluated participants' expectations on when the D-ATCO system should initiate a Request Delegation. Participants indicated conditions for Request Delegation, such as always when capable, only in specific situations (e.g., conflicts with aircraft with shared allocation), to balance workloads, to limit human operator workload, or never.

Analysis Procedure. The UEQ-S data were analyzed descriptively to summarize the user experience ratings. We recorded negatively worded items so that higher values consistently indicate a more positive user experience. Mean scores

were calculated separately for PQ and HQ, and an overall UEQ-S score was computed as the mean of all eight items. Additionally, standard deviations, minimum and maximum values, and 95% confidence intervals were reported for each subscale. The analysis was performed using an Excel Sheet reporting on descriptives, scale consistancy and inconsistance and benchmarks. The self administered questionnaires were analyzed descriptively in Excel.

Focus Group Discussion. The audio data was obtained from a recorded discussion involving ATCOs on the delegation prototype. The recording captured natural, unscripted dialogue where ATCOs provided feedback on the system's usability, benefits, and potential challenges. To facilitate qualitative analysis, the audio data was transcribed into text. This step ensures that spoken content is documented in a written format, making it easier to extract relevant insights. The Otter.ai Transcription tool was used to convert the audio into text, with attention paid to accurately capturing key statements, interruptions, and emphasis. The analysis of the audio data from the qualitative interviews followed a pragmatic qualitative approach, focusing on the extraction of key insights relevant to the evaluation of the delegation prototype. Rather than conducting a full thematic analysis or grounded theory approach, we employed a content reduction strategy to identify and summarize the most salient points from the transcript. This method aligns with qualitative content analysis and rapid assessment techniques, where the goal is to efficiently distill meaningful feedback without an exhaustive coding process.

4 Results

4.1 User Experience Questionnaire Short Version (UEQ-S)

The results from the UEQ-S indicate an overall positive user experience across both PQ and HQ dimensions, with some variation between items. PQ, which reflects the system's usability and efficiency, received consistently positive ratings. All four items in this category scored above 1.4, with the highest ratings for *obstructive - supportive* ($M = 1.9$, $SD = 0.4$) and *complicated - easy* ($M = 1.9$, $SD = 0.4$). These results suggest that participants found the system supportive and easy to use. The lowest score within this dimension was for *inefficient - efficient* ($M = 1.4$, $SD = 1.0$), which, while still positive, indicates some variation in perceived efficiency among participants. The results showed more variation in HQ, which measures the system's aesthetic appeal and innovativeness. Two items-*not interesting - interesting* ($M = 1.7$, $SD = 1.1$) and *usual - leading edge* ($M = 1.6$, $SD = 1.0$)-received positive evaluations, suggesting that users generally found the system engaging and novel. However, the items *boring - exciting* ($M = 0.7$, $SD = 1.0$) and *conventional - inventive* ($M = 0.6$, $SD = 0.8$) fell within the neutral range (values between -0.8 and 0.8), indicating that participants did not strongly perceive the system as particularly exciting or original. Results suggest that the system is functional and enjoyable but it

does not firmly stand out in terms of innovation or excitement. Based on the UEQ-S interpretation framework [16], scores above 0.8 are considered positive, meaning the system was well-received. Given our results, the PQ dimension was uniformly positive. Users thus perceived the system as efficient, easy to use, and supportive. The HQ dimension also received positive feedback, but two items remained neutral, indicating room for improvement in making the system more engaging and innovative.

4.2 User Feedback Delegation

Delegation Function Preferences. Participants indicated their preferred method for initiating Delegation. The majority ($n = 4$, 57.1%) preferred a single mouse click, while one participant ($n = 1$, 14.3%) chose a double mouse click, a separate pop-up confirmation window, and left question unanswered. No participants selected the click-and-hold mouse button.

General Preference for Task Delegation. Most participants supported delegation either for individual aircraft assignments ($n = 3$, 42.9%) or specific tasks for all aircraft within their AOR ($n = 2$, 28.6%). Two participants ($n = 2$, 28.6%) were undecided or only supported task-wise delegation as an additional option, and one participant ($n = 1$, 14.3%) opposed task-wise delegation entirely.

Completion Feedback for Delegation. Participants were asked what feedback to provide upon Delegation completion. The most frequent suggestions included: changing the head symbol ($n = 4$, 57.1%), adding more symbols to the label ($n = 3$, 42.9%), and hiding label information ($n = 3$, 42.9%). Only 14.3% ($n = 1$) suggested relying on visual cues like color, font size, or style changes.

Trigger Situations for Delegation. Participants identified scenarios in which they would personally use Delegation. The most common triggers were open routine tasks that the system (D-ATCO) could handle ($n = 6$, 85.7%) and delegation after resolving a conflict for an aircraft that was only passing through a sector ($n = 6$, 85.7%). Delegation after responding to a pilot request was reported by two participants ($n = 2$, 28.6%), while delegation in overload situations was selected by only one participant ($n = 1$, 14.3%). No participants indicated they would never use Delegation.

4.3 User Feedback Redelegation

Need for Radiocheck After Redelegation. Most respondents ($n = 6$, 85.7%) indicated no need for radio check after Redelegation, while one participant ($n = 1$, 14.3%) was undecided. No participants selected yes.

Priority of Redelegation. Participants ranked different aircraft conditions regarding Redelegation priority, where 1 represented the highest urgency. Aircraft in conflict received the highest priority ($M = 1.14$, $SD = 0.38$), followed by aircraft yet outside their sector but in conflict within the AOR ($M = 2.14$, $SD = 0.69$). Aircraft with coordination and pilot requests received lower priority ratings, with an average rank of $M = 3.14$. However, the latter had a higher standard deviation ($SD = 1.21$), indicating more response variability. One participant specified emergencies as an additional high-priority category.

Trigger Situations for Redelegation. Participants identified situations where they expected the system (D-ATCO) to automatically redelegate control without requiring a request. The most common triggers were conflicts where D-ATCO struggles to find a solution ($n = 7$, 100%) and emergencies ($n = 7$, 100%). Other frequently selected triggers included conflicts with shared allocation between humans and the system ($n = 5$, 71.4%) and weather-affected aircraft ($n = 5$, 71.4%). Coordination tasks ($n = 2$, 28.6%) and pilot requests ($n = 1$, 14.3%) were less frequently considered reasons for Redelegation. One participant specified system failure as an additional condition.

4.4 User Feedback Cancel Delegation

Need for Information to Enable Cancel Delegation. All participants ($n = 7$, 100%) indicated they need no information to enable cancellation of delegation, suggesting that they did not see a need for additional confirmation before canceling a delegated task.

Required Information to Confirm Cancel Delegation Completion. Participants could select multiple options regarding how they preferred to observe the completion of a canceled delegation. The most commonly chosen feedback mechanism was a change in the head symbol ($n = 4$, 57.1%). Other preferred methods included extension of label information ($n = 3$, 42.9%), additional symbols/characters in the label ($n = 2$, 28.6%), and changes in label color, font size, or style ($n = 2$, 28.6%). No participants selected attention guidance (e.g., highlighting the head symbol with a filled colored circle).

Need for Radiocheck After Cancel Delegation. All participants ($n = 7$, 100%) indicated that a radiocheck was not necessary after Cancel Delegation.

4.5 User Feedback Request Delegation

Conditions for D-ATCO to Request Delegation. Participants were asked under which conditions the system (D-ATCO) should Request Delegation. The majority ($n = 6$, 85.7%) preferred that D-ATCO should always Request Delegation if it is able to handle a flight. Additionally, three participants (42.9%) supported delegation to balance workload between humans and D-ATCO, while the same number of participants ($n = 3$, 42.9%) agreed with delegation to cap the human workload. No participants selected never.

4.6 Semi-Structured Interview

Analysing the transcripts from the qualitative discussions on the delegation prototype showed that theme that usability and interaction design was frequently mentioned as key point. ATCOs expressed a preference for minimal interaction steps, indicating that even two clicks for delegation might feel excessive. Keeping label information visible at all times to maintain situational awareness. Critiques were directed at the Human-Machine Interface (HMI), particularly regarding the complexity of interactions for routine tasks was also considered as highly relevant. At times, the display of flight levels was unclear, causing confusion about actual clearance levels. To effectively communicate delegation status and avoid interface clutter, ATCOs suggested to consider using symbols or slight modifications to labels. They also favored interactions through radar labels over separate menu options, as the former allows for more direct control.

ATCOs also highlighted the need for seamless delegation. They emphasized that the system should clearly communicate why and when Redelegation occurs, as some found themselves only noticing it after closely examining the display. There was also concern about whether the system sufficiently considered real operational complexities, such as coordination with adjacent sectors and handling unexpected pilot requests.

Cognitive load and situational awareness were further discussed, with some ATCOs acknowledging a steep learning curve with the system, stating that it was not intuitively designed. ATCOs indicated that they would like to have full control over all operations. Some of them were even concerned that Redelegation might be conducted without clear justification. Suggestions included incorporating small tooltips or hover-over explanations to enhance understanding without causing distractions.

When it came to system intelligence and decision-making, ATCOs had expectations that the system should anticipate conflicts and notify them in advance, especially regarding expected coordination issues with adjacent sectors, pilot expectations versus actual system behavior, and situations where Redelegation is unavoidable. ATCOs were also concerned about automation bias. They wondered whether the system would consistently make correct decisions and preferred manual interventions in edge cases. Therefore it would be necessary to easily override automation and ensure that the system does not assume full control without their consent.

In terms of recommendations for improvement, ATCOs urged reducing the number of clicks required for delegation and redelegation to streamline the process. They insisted on keeping label information visible at all times to prevent a loss of situational awareness and improving system transparency by providing clear indications of why a delegation or redelegation occurred. As essential were considered to display information directly in the radar label to minimize the need for additional screens and pop-ups. Additionally, ATCOs considered confirmation steps only as necessary if they serve operational purposes. ATCOs furthermore clearly requested better visual cues like subtle color changes or icons to better differentiate delegation statuses. Tooltips or hover-over expla-

nations should assist them in understanding the decisions of D-ATCO without increasing their cognitive load. Finally, they required that the system considers or detects their personal preferences regarding the timing and manner of delegation decisions, particularly in relation to coordination with other sectors.

5 Discussion and Conclusion

This poster provides a comprehensive analysis of the feedback and recommendations received from ATCOs during the evaluation of the delegation process with a D-ATCO prototype. The ATCOs provided valuable insights into the usability, interaction design, and system intelligence of the prototype.

Our study's findings are limited by the predefined scenarios used for Delegation, Redelegation, and Request Delegation, which may not accurately represent real-world ATC environments due to their narrow scope. Additionally, the D-ATCO system was not fully functional and required pre-allocated aircraft to simulate delegation processes, which introduces an artificial constraint. The relatively small number of ATCOs involved in our study also limits the statistical power of our results. It's also possible that some usability aspects were not thoroughly explored during discussions, potentially affecting the validity of our findings.

To build upon our research, it is envisaged to test delegation functionalities in real-time ATC simulations with dynamic traffic and realistic sector constraints. This can help to assess whether the delegation approach aligns with ATCO's operational decision-making. Moreover, it's crucial to investigate how trust in the system evolves over prolonged use to ensure its long-term effectiveness.

Acknowledgments. The authors would like to extent their sincere gratitude to Carmen Bruder (DLR) and Jörg Buxbaum (DFS) for their outstanding efforts in organizing this successful workshop. Their tireless work and adaptability in the face of unforeseen circumstances were truly appreciated. Furthermore, we like to thank all participating ATCOs from DFS and Austrocontrol, who joined us with enthusiasm and expertise.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bruder, C., Theis, S., Stefani, T., Jameel, M., Hunger, R., Schulze Kissing, D.: User participation in the design of trustworthy human-AI-collaboration in air-traffic control (06 2024). https://doi.org/10.13140/RG.2.2.27418.25287
2. DLR Homepage: https://www.dlr.de/en/fl/research-transfer/projects/loki. Accessed 07 Mar 2025
3. Endsley, M.R.: From here to autonomy. Hum. Factors **59**(1), 5–27 (2017). https://doi.org/10.1177/0018720816681350
4. EUROCONTROL: Aircraft Performance Database. https://contentzone.eurocontrol.int/aircraftperformance/details.aspx?ICAO=A20N. Accessed 10 Mar 2025

5. EUROCONTROL: EUROCONTROL Forecast 2024–2030. https://www.eurocontrol.int/publication/eurocontrol-forecast-2024-2030. Accessed 07 Mar 2025
6. EUROCONTROL: PRR 2023 Peformance Review Report. https://www.eurocontrol.int/sites/default/files/2024-06/eurocontrol-performance-review-report-2023.pdf. Accessed 26 Feb 2025
7. Garbis, C., Artman, H.: Team situation awareness as communicative practices. A Cognitive Approach to Situation Awareness: Theory and Application, pp. 275–296 (2004)
8. Gerdes, I., Jameel, M., Hunger, R., Christoffels, L., Gürlük, H.: the automation evolves: concept for a highly automated controller working position. In: 33rd Congress of the International Council of the Aeronautical Sciences. ICAS, Stockholm (2022)
9. Hendrickx, P., Tisza, A.: EUROCONTROL Maastricht Upper Area Control Centre OPS & Automation Strategy. https://skybrary.aero/sites/default/files/bookshelf/5341.pdf. Accessed 26 Feb 2025
10. Jameel, M., Tyburzy, L., Gerdes, I., Pick, A., Hunger, R., Christoffels, L.: Enabling digital air traffic controller assistant through human-autonomy teaming design. In: DASC 2023 conference proceedings. pp. 1–9. IEEE, Piscataway, NJ, USA (2023). https://doi.org/10.1109/DASC58513.2023.10311220
11. Kühn, M.: Fuel consumption of the 50 most used passenger aircraft. https://doi.org/10.7910/DVN/4CYNKA
12. Laugwitz, B., Held, T., Schrepp, M.: Construction and Evaluation of a User Experience Questionnaire. In: Holzinger, A. (ed.) HCI and usability for education and work, pp. 63–76. Lecture Notes in Computer Science, Springer, Berlin and Heidelberg (2008). https://doi.org/10.1007/978-3-540-89350-9_6, https://link.springer.com/chapter/10.1007/978-3-540-89350-9_6
13. Rogosic, T., Juricic, B., Antulov-Fantulin, B., Radisic, T.: Review of automation problems in air traffic control. In: Koricic, M. (ed.) 44th International Convention on Information, Communication and Electronic Technology (MIPRO), pp. 1636–1641. IEEE, Piscataway, NJ, USA (2021). https://doi.org/10.23919/MIPRO52101.2021.9597063
14. de Rooij, G., Borst, C., Van Paassen, M.M., Mulder, M.: Flight allocation in shared human-automation en-route air traffic control. In: 21st International Symposium on Aviation Psychology, pp. 172–177 (2021)
15. de Rooij, G., Tisza, A., Borst, C., van Paassen, R., Mulder, M.: Towards human-automation teamwork in shared en-route air traffic control: task analysis. In: Kaber, D. (ed.) Proceedings of the 2022 IEEE International Conference on Human-Machine Systems (ICHMS), pp. 1–7. IEEE, Piscataway, NJ, USA (2022). https://doi.org/10.1109/ICHMS56717.2022.9980715
16. Schrepp, M., Hinderks, A., Thomaschewski, J.: Design and evaluation of a short version of the user experience questionnaire (UEQ-S). Int. J. Interact. Multimed. Artif. Intell. **4**, 103 (2017). https://doi.org/10.9781/ijimai.2017.09.001
17. Shively, R.J., Lachter, J., Brandt, S.L., Matessa, M., Battiste, V., Johnson, W.W.: Why Human-Autonomy Teaming? In: Baldwin, C. (ed.) Advances in Neuroergonomics and Cognitive Engineering AHFE 2017, vol. 586, pp. 3–11. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-60642-2_1

Conceptual Model of Information Sharing and Interaction in Role-Sharing for High Team Performance Through a Pilot Study

Jun Ichikawa[1(✉)], Masatoshi Yamada[2(✉)], Yutaka Iwaihara[3], and Genki Ichinose[1]

[1] Shizuoka University, 3-5-1 Johoku, Chuo-ku, Hamamatsu, Shizuoka, Japan
j-ichikawa@inf.shizuoka.ac.jp, ichinose.genki@shizuoka.ac.jp
[2] Tokoha University, 6-1 Yayoi-cho, Suruga-ku, Shizuoka, Shizuoka, Japan
yamada-m@sz.tokoha-u.ac.jp
[3] Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, Shizuoka, Japan
iwaihara.yutaka@shizuoka.ac.jp

Abstract. This study aims to understand complex and dynamic group coordination, wherein three or more members interact over time through role-sharing to achieve a common goal. We proposed a conceptual model of information sharing and interaction in role-sharing for high team performance. This model will guide designs for coordination between humans and autonomous AI. As a pilot study, we analyzed interview data recorded in a field experiment observing offensive coordination in 3-on-3 basketball. Six top-level female players from Shizuoka Prefecture who competed at the National Sports Festival of Japan were divided into offensive and defensive teams. A mini-game designed in a previous study was repeated. The offensive players explained crucial interactions, shared ideas, and anticipated teammates' plays, etc. The constructive concepts of offensive strategies were extracted from each player's report, and their co-occurrence relationships were networked. The results confirmed shared and individual-specific concepts, respectively. The shared concept was prominent across the players. The individual-specific concept was prominent for each player. The shared concepts acted as hubs and tended to have higher co-occurrence relationships with some individual-specific concepts. Considering quantitative and qualitative results, we propose a conceptual model comprising three elements: (A) sharing a core strategy related to a mental model for efficient interaction, (B) leaving to other roles without information sharing, which may trigger flexible interaction, and (C) relying on implicit anticipation and trust to support (A) and (B). This study provides fundamental findings for understanding complex and dynamic coordination. Future studies should analyze group behavior to investigate the three elements.

Keywords: Group coordination · Role-sharing · Information sharing · Conceptual model

J. Ichikawa and M. Yamada—Equal contribution.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 323–331, 2025.
https://doi.org/10.1007/978-3-031-94171-9_31

1 Introduction

Humans can share roles with other members in various social activities. Role-sharing assumes that it is preferable to leave to others for achieving a common goal, considering the quality and quantity of activities to be done. In cognitive science, which discusses human information processing underlying behaviors, interactions with others to achieve a group goal are called planned coordination [1]. This study investigated information sharing and interaction in role-sharing for high team performance.

As related work, coordinated role-sharing has been investigated across research fields, such as cognitive science, sports science, and network science. Cognitive science discusses the necessity of establishing a mental model as a representation of knowledge structure based on top-down information processing. Such a shared or complemented structure among members leads to efficient interaction (e.g., [2,3]). It is suggested that member changes carry the risk of collapsing the mental model and reducing team performance [4]. Moreover, achieving a group goal requires integrating others' perspectives into one's own through role-sharing and -alternating. (e.g., [5,6]). Thus, for high performance, information sharing of knowledge and perspectives, and cognitive interaction among members are vital. However, considering the qualities and quantities of activities and their loads, sharing all information is not feasible. Establishing a transactive memory system is also crucial, which is some parts of the activities are left to others, and related information is distributed (e.g., [7,8]). Menwhile, sports science and network science investigate coordinative physical structures in team sports, such as soccer and basketball. The sports science study recorded a 5-on-5 basketball game by a top-level university team in Japan and confirmed diverse role-sharing of defensive coordination depending on emergent levels [9]. The network science study explained the characteristics related to robust passing with role-sharing by network models, using player tracking data recorded in professional soccer games (e.g., [10,11]). These findings suggest that expert teams can handle various situations flexibly.

According to the related work and discussion above, the degree of information sharing must be considered to understand coordinated role-sharing. Establishing a mental model based on information sharing enables efficient interaction. Meanwhile, leaving to other members and interacting among roles without infromation sharing may be also required to trigger flexible handling. However, information processing underlying complex and dynamic coordination has not been fully investigated, wherein three or more members interact over time through diverse relationships in role-sharing [12,13]. Cognitive science focuses on human processing and often uses controlled experimental tasks to investigate relationships between factors and performance directly. Few studies have used complex and dynamic tasks, such as team sports. Sports science and network science analyze complex and dynamic physical structures in team sports; however, the information processing underlying these characteristics has not been fully discussed.

We aim to understand complex and dynamic coordination. Therefore, our approach was to observe role-sharing in real-world. The purpose of this study

was to propose a conceptual model of information sharing and interaction in role-sharing for high team performance. As a pilot study, we conducted a field experiment, wherein the experts played a 3-on-3 basketball mini-game. At the first step in this challenging study, we analyzed interview data on offensive coordination. The proposed model provides fundamental findings and contributes to future designs for coordination between humans and autonomous AI.

2 Methods

2.1 Participants

Six top-level female players from Shizuoka Prefecture who competed at the National Sports Festival of Japan, known as JAPAN GAMES, participated in the field experiment. This study was approved by the ethics and safety committee. Written informed consent was obtained from all the players.

2.2 Procedures

In the field experiment, the six players were divided into the offensive and defensive teams. Figure 1 illustrates the 3-on-3 basketball mini-game with a 15-s time limit, designed in a previous study [13] to investigate offensive coordination. As a winning condition, the offensive team must ensure that a player draws a

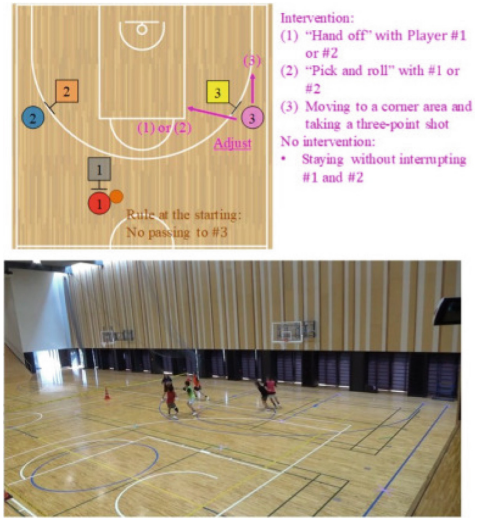


Fig. 1. Diagram of the 3-on-3 basketball game and field experiment. The upper portion illustrates the starting position for each player in the game, designed in the previous study [13]. This is cited under CC-BY. The lower portion shows the image recorded by a video camera. The circle and square represent the offensive and defensive players, respectively, with each color indicating the worn bib.

defensive player and a shooter is unmarked in open space. Some rules were set in this experiment: the starting position of each offensive player was fixed, and the defensive players matched against them as 1-on-1 were also fixed. Additionally, the ball handler (#1 red) was fixed, and the game began without a direct pass to #3 pink, who was the furthest from #1 red. Three sessions of seven trials per session were conducted for 21 trials. To win the game for the offensive team, #3 pink is a key player; this role needs to support the other offensive roles: one who dribbles the ball (#1 red) and another who interacts with #1 red to create a shooting chance near the goal (#2 blue). Examples include (1) directly handing the ball off, known as "hand off"; (2) acting as a wall to interrupt the defensive players, known as "pick and roll"; and (3) moving to a corner area and taking a three-point shot. Based on the situations, it is also crucial not to interrupt plays by #1 red and #2 blue. These fundamental offensive plays are essential and applicable to a 5-on-5 game. The details have been explained in the previous study[1].

After repeating the game in the field experiment, the interview was conducted. The offensive team was asked to answer six questions: (1) crucial interactions, (2) shared ideas related to strategies for actualizing (1), (3) anticipated teammates' plays for actualizing (1), (4) shared ideas related to strategies for defense, (5) anticipated defensive movements, and (6) anticipated defense based on one's own offensive interactions, respectively. Each player was given opportunities to explain them.

2.3 Analysis

Constructive concepts were extracted from recorded reports by the Steps for Coding and Theorization (SCAT) used in qualitative research. The procedures included (1) picking up noteworthy words or phrases from the text, (2) paraphrasing of (1), (3) conceptualizing from out of the text that account for (2), and (4) emerging themes and constructs in considerations of context. These referred to the tutorials papers [14,15], and a previous study [16] that applied SCAT to sports players' reports. Furthermore, the concepts were checked and corrected by the third author, who was the head coach, to increase qualitative validity. The ratio of each concept type to the total number of concepts was calculated for each player.

Additionally, co-occurrence relationships among the extracted concepts were networked for each player using the dialogue analysis tool Knowledge Building Discourse eXplorer (KBDeX) [17].

3 Results and Discussion

Table 1 represents the constructive concepts of offensive strategies extracted using SCAT. A total of 10 concepts are listed, with the top two ratios, including

[1] The original description of the mini-game refers to the previous study with CC-BY [13].

ties, highlighted in bold. These are (1) space to play, (2) interaction related to the starting positions, as shown in Fig. 1, (3) interaction with a teammate on the opposite side of the ball, (4) interaction related to strategies for defense on the opposite side of the ball, (5) defense itself on the opposite side of the ball, (6) personally excelled plays and roles, (7) understanding of teammates' plays related to concept (6), (8) interaction which is challenging to explain, (9) defense itself which is challenging to explain, and (10) ones other than them. The ratio of concept (3) was higher than those of the others across all the players.

Additionally, the ratios of some constructive concepts were prominent for each player: concepts (2), (8), and (9) for #1 red, concept (6) for #2 blue, and concept (5) for #3 pink. Meanwhile, the shared and individual-specific concepts related to individual plays and interactions (#1 red: concept (2); #2 blue: concept (6); #3 pink concept (5)) tended to be extracted with concepts (7), (8), and (9), which were fundamental for movements. The qualitative reports suggested that the shared and individual-specific concepts were supported by implicit anticipation and trust (e.g., "For coordination, I can understand #3 pink's position through sufficient experience without watching, as the play to create open space is a common strategy across teams. There is a common consensus without explicit instructions, similar to dribbling to the goal and passing. It is challenging to explain them."[2]).

Table 1. Observed constructive concepts of offensive strategies. The ratio of each concept type to the total number of concepts is calculated for each player. The top two ones, including ties, are highlighted in bold.

No.	Constructive concept	Players		
		#1 red	#2 blue	#3 pink
(1)	Space	0.053	0.083	0.100
(2)	Interaction of a 2-man game	**0.158**	0.125	0.150
(3)	Interaction with a teammate on the help side	**0.211**	**0.167**	**0.250**
(4)	Interaction related to strategies for defense on the help side	0.053	0.042	0.000
(5)	Defense	0.053	0.125	**0.250**
(6)	Own play and role	0.053	**0.167**	0.050
(7)	Understanding of teammates' plays	0.105	0.083	0.150
(8)	Implicit knowledge of interaction	**0.158**	0.083	0.050
(9)	Implicit knowledge of defense	**0.158**	0.000	0.000
(10)	Others	0.000	0.125	0.000

Note. Constructive concepts were extracted from the explanations in the interviews of the field experiment. These comprised the following 10 types: (1) space to play; (2) interaction related to the starting positions, as shown in Fig. 1; (3) interaction with a teammate on the opposite side of the ball; (4) interaction related to strategies for defense on the opposite side of the ball; (5) defense itself on the opposite side of the ball; (6) personally excelled plays and roles; (7) understanding of teammates' plays related to concept (6); (8) interaction which is challenging to explain; (9) defense itself which is challenging to explain, and (10) ones other than them.

[2] Concept (3), (6), (7), and (8) were extracted from the #1 red's report.

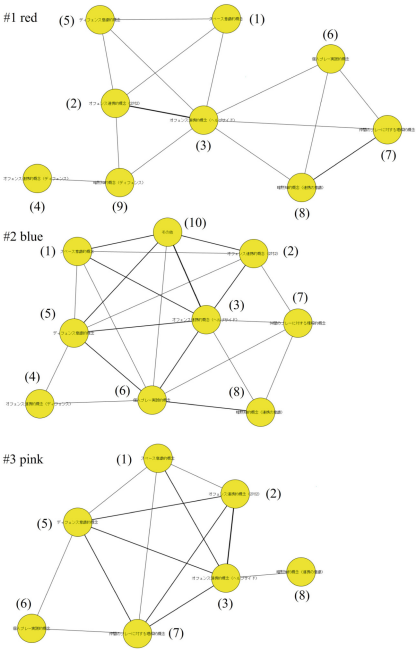


Fig. 2. Co-occurrence networks showing the relationships among the constructive concepts of offensive strategies. Each node represents a concept, as shown in Table 1. The thicker the link, the higher the co-occurrence. For all the players, the concept of (3) interaction with a teammate on the help side acts as a hub. This shared concept across the players tends to have higher co-occurrence relationships with some individual-specific concepts, such as (2) interaction of a 2-man game by #1 red, (6) own play and role by #2 blue, and (5) defense by #3 pink, respectively. These are prominent for each player and shown in bold in Table 1.

Next, Fig. 2 illustrates co-occurrence networks showing the relationships among the constructive concepts of offensive strategies, using KBDeX. Each node represents a concept, as listed in Table 1. The thicker the link, the higher the co-occurrence. Shared concept (3) across all the players acted as hubs. Table 2 represents the degree and closeness centralities for each player. The degree centrality indicates the extent to which a node is directly connected to others, while the closeness centrality measures how many steps are required to reach other nodes from a particular node. The values of concept (3) for both degree and closeness centralities were higher than those of the others. Concept (3) tended to have higher co-occurrence relationships with some individual-specific concepts, such as concept (2) for #1 red, concept (6) for #2 blue, and concept (5) for #3 pink, all of which were indicated individual plays and interactions as mentioned above.

Considering the quantitative and qualitative results, we propose a conceptual model of information sharing and interaction in role-sharing for high team

Table 2. Degree and closeness centralities of constructive concepts related to offensive strategies in the co-occurrence networks, as shown in Fig. 2. The values highlighted in bold represent the highest and correspond to the shared concept in Table 1.

No.	Constructive concept	Players					
		#1 red		#2 blue		#3 pink	
		Degree centrality	Closeness centrality	Degree centrality	Closeness centrality	Degree centrality	Closeness centrality
(1)	Space	0.375	0.571	0.625	0.727	0.667	0.750
(2)	Interaction of a 2-man game	0.500	0.667	0.625	0.727	0.667	0.750
(3)	Interaction with a teammate on the help side	**0.875**	**0.889**	**0.875**	**0.889**	**0.833**	**0.857**
(4)	Interaction related to strategies for defense on the help side	0.125	0.400	0.250	0.571	0.000	0.000
(5)	Defense	0.375	0.571	0.750	0.800	0.833	0.857
4 (6)	Own play and role	0.375	0.571	0.875	0.889	0.333	0.545
(7)	Understanding of teammates' plays	0.375	0.571	0.500	0.667	0.833	0.857
(8)	Implicit knowledge of interaction	0.375	0.571	0.375	0.615	0.167	0.500
(9)	Implicit knowledge of defense	0.375	0.615	0.000	0.000	0.000	0.000
(10)	Others	0.000	0.000	0.625	0.727	0.000	0.000

Note. Degree centrality indicates the extent to which a node is directly connected to others. Closeness centrality measures how many steps are required to reach other nodes from a particular node.

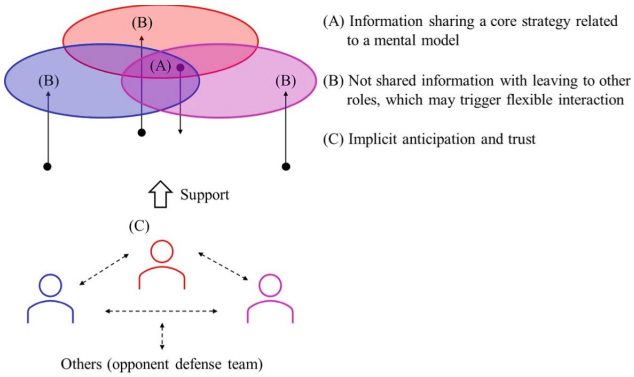


Fig. 3. Conceptual model of information sharing and interaction in role-sharing for high team performance. Three elements act: (A) sharing a core strategy related to a mental model for efficient interaction, (B) leaving to other roles without information sharing, which may trigger flexible interaction, and (C) relying on implicit anticipation and trust to support (A) and (B).

performance (Fig. 3). This comprises three elements: (A) sharing a core strategy related to a mental model for efficient interaction, (B) leaving to other roles

without information sharing as a transactive memory system, which may trigger flexible interaction, and (C) relying on implicit anticipation and trust to support (A) and (B). This model is consistent with the findings of previous studies (e.g. [3,8]). After the first and second authors designed the model, the third author provided feedback for future investigation. The proposed model explains actual plays well, regardless of the field experiment. It could be improved by incorporating a core strategy shared by the defense team into (A). This study provides fundamental findings for understanding complex and dynamic coordination. Future studies should analyze group behavior to investigate the three elements.

4 Conclusion

This study proposed a conceptual model of information sharing and interaction in role-sharing for high team performance. As a pilot study, we conducted the field experiment wherein the experts played the 3-on-3 basketball mini-game. The interview results of offensive coordination through exploratory analysis showed that the conceptual model comprises three elements: (A) sharing a core strategy related to a mental model for efficient interaction, (B) leaving to other roles without information sharing, which may trigger flexible interaction, and (C) relying on implicit anticipation and trust to support (A) and (B). Our model will contribute to designs for complex and dynamic coordination between humans and autonomous AI. For example, they do not need to share all information in social activities, implying that only a core strategy is shared. It is important to leave to each other through role-sharing. Additionally, as a prerequisite, they should mutually enhance trust and the accuracy of various anticipation. Future work should analyze group behavior to investigate the validity of this model.

Acknowledgments. We would like to thank Yugo Takeuchi (Shizuoka University) for meaningful comments. This study was supported by JSPS KAKENHI Grant Number 24K20562, Daiichi-Sankyo "Habataku" Support Program for the Next Generation of Researchers, AY 2024 Grant of Interdisciplinary Research Promotion by Headquarters for Promotion of Interdisciplinary Domain Research in Shizuoka University, and AY 2024 Research Grant of Establishment of Interdisciplinary Networks for Development of Innovative Areas in Shizuoka University.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Knoblich, G., Butterfill, S., Sebanz, N.: Psychological research on joint action: Theory and data. In: Ross, B. (ed.) The Psychology of Learning and Motivation: Advances in Research and Theory, vol. 54, pp. 59–101. Academic Press, Cambridge (2011)

2. Rico, R., Sánchez-Manzanares, M., Gil, F., Gibson, C.: Team implicit coordination processes: a team knowledge-based approach. Acad. Manag. Rev. **33**(1), 163–184 (2008)
3. Steiner, S., Macquet, A.C., Seiler, R.: An integrative perspective on interpersonal coordination in interactive team sports. Front. Psychol. **8**, 1440 (2017). https://doi.org/10.3389/fpsyg.2017.01440
4. Bedwell, W.L.: Adaptive team performance: The influence of membership fluidity on shared team cognition. Front. Psychol. **10**, 2266 (2019). https://doi.org/10.3389/fpsyg.2019.02266
5. Shirouzu, H., Miyake, N., Masukawa, H.: Cognitively active externalization for situated reflection. Cogn. Sci. **26**(4), 469–501 (2002)
6. Hayashi, Y.: The power of a "Maverick" in collaborative problem solving: an experimental investigation of individual perspective-taking within a group. Cogn. Sci. **42**, 69–104 (2018)
7. Hollingshead, A.B.: Cognitive interdependence and convergent expectations in transactive memory. J. Person. Social Psychol. **81**(6), 1080–1089 (2001). https://doi.org/10.1037/0022-3514.81.6.1080
8. Lewis, K., Belliveau, M., Herndon, B., Keller, J.: Group cognition, membership change, and performance: Investigating the benefits and detriments of collective knowledge. Organ. Behav. Human Decision Process. **103**(2), 159–178 (2007). https://doi.org/10.1016/j.obhdp.2007.01.005
9. Fujii, K., Yokoyama, K., Koyama, T., Rikukawa, A., Yamada, H., Yamamoto, Y.: Resilient help to switch and overlap hierarchical subsystems in a small human group. Sci. Reports **6**, 23911 (2016). https://doi.org/10.1038/srep23911
10. Buldú, J.M., Busquets, J., Echegoyen, I., Seirul.lo, F.: Defining a historic football team: Using Network Science to analyze Guardiola's F.C. Barcelona. Sci. Reports **9**, 13602 (2019). https://doi.org/10.1038/s41598-019-49969-2
11. Ichinose, G., Tsuchiya, T., Watanabe, S.: Robustness of football passing networks against continuous node and link removals. Chaos, Solitons Fractals **147** (2021). https://doi.org/10.1016/j.chaos.2021.110973
12. Ichikawa, J., Fujii, K.: Proposal of a research approach for discussion of a dynamic coordination mechanism: investigation of anticipating others' behaviors and adaptation through quantitative analysis of group behavior. Cogn. Stud.: Bull. Japanese Cogn. Sci. Society **27**(3), pp. 377–385 (2020) https://doi.org/10.11225/cs.2020.026
13. Ichikawa, J., Yamada, M., Fujii, K.: Analysis of coordinated group behavior based on role-sharing: Practical application from an experimental task to a 3-on-3 basketball game as a pilot study. bioRxiv (2024) https://doi.org/10.1101/2024.09.16.612561
14. Otani, T.: SCAT a qualitative data analysis method by four-step coding: easy startable and small scale data-applicable process of theorization. Bull. Graduate School Educ. Human Develop. Nagoya Univ. **54**(2), 27–44 (2008)
15. Otani, T.: SCAT: Steps for Coding and Theorization. J. Japan Society of Kansei Eng. **10**(3), 155–160 (2011). https://doi.org/10.5057/kansei.10.3_155
16. Yamada, M., Ohata, M., Kakoi, D.: A study on cognitive transformation in the process of acquiring movement skills for changing running direction. IEICE Trans. Inform. Syst. **105**(3), 565–577 (2022). https://doi.org/10.1587/transinf.2021HCP0007
17. Oshima, J., Oshima, R., Matsuzawa, Y.: Knowledge building discourse explorer: a social network analysis application for knowledge building discourse. Educ. Technol. Res. Develop. **60**(5), 903–921 (2012). https://doi.org/10.1007/s11423-012-9265-2

Understanding Human-AI Collaboration: A Systematic Review of Challenges and Research Methods in Management

Rafał Łabędzki(✉), Katarzyna Mikołajczyk, Anna Biłyk, and Monika Trojanowska

SGH Warsaw School of Economics, al. Niepodległości 162, 02-554 Warszawa, Polska
{rlabed,katarzyna.mikolajczyk,anna.bilyk}@sgh.waw.pl, monika.trojanowska@onet.pl

Abstract. Research on collaboration between human and non-human intelligent agents is extensive, yet a deeper understanding of the specific management and organizational challenges and the research methods used to study them is still needed. A systematic literature review was conducted following methodological guidelines to address this gap, with the PRISMA 2020 flow diagram used to document the process. This study explores two key research questions: (RQ1) What specific management and organizational challenges are investigated in human-AI collaboration? (RQ2) What research methods are commonly employed in this field? Data were retrieved from Scopus (486 documents) and Web of Science (385 documents), resulting in 95 studies in the final analysis. The review identifies five key categories of management and organizational challenges: (1) Hybrid human-AI teams, (2) AI integration in work processes, (3) Human-AI decision-making, (4) Trust in AI, and (5) Human-robot collaboration (HRC). Quantitative methods were more frequently used than qualitative approaches, with experimental studies—particularly simulation designs—being the most common. In qualitative research, grounded theory and case study designs were equally prominent. The findings contribute to the ongoing discourse on human-AI collaboration by mapping key challenges and methodological trends, providing insights for future research and organizational practice .

Keywords: Human-AI Collaboration · Human-Robot Collaboration · Hybrid Multi-Agent Systems · Management

1 Introduction

With growing interest in Hybrid Multi-Agent Systems (HyMAS) in management research, understanding the challenges of human-AI collaboration is becoming increasingly important. While HyMAS remains conceptual, various studies have explored cooperation between human and artificial agents (A. R. Dennis et al., 2023; G. Zhang et al., 2023). However, the specific challenges studied and the methods used to examine them remain unclear. A systematic literature review can provide valuable insights into the current state of research and the methodologies applied in this field.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 332–348, 2025.
https://doi.org/10.1007/978-3-031-94171-9_32

Although no prior systematic review directly addresses our research questions, multiple reviews explore different dimensions of human-AI interaction (Bilan et al., 2022; Gebru et al., 2022; Madanchian et al., 2023; Taboada et al., 2023; Votto et al., 2021). Management research has primarily investigated human-AI collaboration within the contexts of human resource management and project management (Madanchian et al., 2023; Prikshat et al., 2023; Qamar et al., 2021; Seeber et al., 2020; Taboada et al., 2023; Votto et al., 2021). However, existing reviews tend to focus on management tools and processes, rather than the management and organizational challenges or the research methods used in this area.

This review addresses this gap by systematically categorizing the management and organizational challenges in human-AI collaboration and identifying the field's most commonly employed research methods.

2 Method

A systematic literature review was selected as the research method (Sutton et al., 2019). The study followed the standard systematic review process, including (1) defining research questions, (2) data extraction, (3) quality assessment, and (4) synthesis (Klatt, 2023). To document the data extraction and quality assessment process, the PRISMA 2020 flow diagram was utilized (Page et al., 2021). The inclusion and exclusion criteria were independently applied by two blind reviewers, with a third reviewer resolving any discrepancies. The study aimed to answer the following research questions:

RQ1: What specific management and organizational challenges are examined in the context of human-AI collaboration?

RQ2: What research methods are most commonly used to study human-AI collaboration?

Since this study does not address a particular organizational problem but rather focuses on the topics explored and research methodologies applied in the field, the CIMO framework was not employed in this review (Klatt, 2023).

Data were obtained from two databases: Scopus (486 documents) and Web of Science (385 documents). These databases were chosen for their ability to filter searches using the pre-defined inclusion criteria. These databases were selected as they only permitted queries based on the established inclusion criteria.

The selection process included articles, conference papers, and book chapters published in English within the fields of business, management, and accounting. Only publications from 2016 onward were considered. This timeframe aligns with the emergence of Hybrid Multi-Agent Systems (HyMAS) leveraging large language models based on transformer architecture (Vaswani et al., 2017) and its successors. During the abstract screening and full-text eligibility assessment, the following inclusion criteria were applied: 1) The study must be relevant to the field of management, 2) The text must present an original empirical study, and 3) The study must examine human-AI collaboration. A single exclusion criterion was applied: the absence of an abstract. The key criterion was the presence of human-AI collaboration. This was determined by applying an operational definition of autonomous artificial intelligent agents (Łabędzki, 2024).

3 Results

The initial search identified 871 records, with 123 duplicates removed, leaving 748 records for abstract screening. After screening the abstracts and applying the inclusion/exclusion criteria, 465 studies were excluded. As a result, 283 texts were selected for retrieval, of which 190 were successfully obtained. Following a full-text screening, 95 studies met the eligibility criteria and were included in the review. The details of the abstract screening process are presented in Table 1.

Table 1. PRISMA flow.

Phase	Records	Records removed
Database query	Records identified from: Databases (n = 871) Registers (n = 0)	Records removed before the screening: Duplicate records removed (n = 123) Records marked as ineligible by automation tools (n = 0) Records removed for other reasons (n = 0)
Abstract screening	Records screened (n = 748)	Records excluded (n = 465)
	Reports sought for retrieval (n = 283)	Reports not retrieved (n = 93)
Full text screening	Reports assessed for eligibility (n = 190)	Reports excluded (one study could be excluded for more than one reason): • text must present original empirical study (n = 10) • study must investigate human-ai collaboration (intelligent agent autonomy) o Human-Robot Collaboration (HRC) with non-autonomous robots (we left either no HRC at all or HRC with autonomous robots) = 81 o no human-ai collaboration (for both no HRC and HRC with autonomous-robots) = 65 • text must be valid for the field of management (n = 0)
	Total studies included in review (n = 95)	

Table 2. Research methods and designs employed in studies of human-AI collaboration[1].

Research Method	Research Design	Number of assigned studies	Assigned studies
Quantitative Methods	Survey Research/Cross-sectional studies	7	(Ameen et al., 2024; Berx et al., 2024; Franken et al., 2023; Hoch et al., 2024; Siemon, 2022; T. Wu et al., 2022; Zywiolek, 2024)
	Survey Research / Longitudinal studies	1	(T.-J. Wu et al., 2024)
	Experimental Research / True experiments	19	(Blaurock et al., 2024; Cheng et al., 2022; Chu et al., 2024; A. Dennis et al., 2023; Fügener et al., 2022; Georganta & Ulfert, 2024a, 2024b; Gnewuch et al., 2023; Gonzalez-Jimenez & Pinto, 2024; Haupt et al., 2024; Jain et al., 2023; Karuppiah et al., 2023; Lehmann et al., 2022; Molitor & Renkema, 2022; Pareek et al., 2024; Pasparakis et al., 2023; Shanks et al., 2024; T.-J. Wu et al., 2024; Xu et al., 2024)
	Experimental Research / Quasi-experiments	15	(Adam et al., 2024; Chand et al., 2023; Ellwart et al., 2022; Fügener et al., 2021; Gaczek et al., 2023; Lin & Lukodono, 2022; Lv et al., 2021; Peng et al., 2024; Pettersson & Falkman, 2024; W. Wang et al., 2024; Xia et al., 2024; Yun & Jun, 2022; Zanchettin & Lacevic, 2022; K. Zhang et al., 2020; M. Zhang et al., 2022a, 2022b)
	Experimental Research / Single-subject designs	0	

(*continued*)

[1] Note. Classification framework is based on (Creswell & Creswell, 2018; Davis et al., 2007; Happach & Tilebein, 2015).

Table 2. (*continued*)

Research Method	Research Design	Number of assigned studies	Assigned studies
	Experimental Research / Simulation	30	(Azadeh et al., 2023; Bouarfa et al., 2018; Cimpeanu et al., 2020; Fan & Zheng, 2024; Ferreira et al., 2021; Hu et al., 2020; Kattel et al., 2020; Kridalukmana et al., 2019; Lin et al., 2022; Listopad, 2020; T. Liu et al., 2023; Malik et al., 2020; Mao et al., 2023, 2024a, 2024b; Mo et al., 2023; Nourmohammadi et al., 2022; Oliff et al., 2020; Rodriguez & Hilaire, 2023; Rumovskaya, 2020; Srinivas & Yu, 2022; Q. Wang et al., 2023; W. Wang et al., 2024; Weckenborg et al., 2020; J. Wu et al., 2023; Yu et al., 2021; Yun & Jun, 2022; R. Zhang et al., 2022a, 2022b; T. Zhang et al., 2024; Zhu et al., 2024)
Qualitative Methods	Narrative Research	0	
	Phenomenological Research	0	
	Grounded Theory	18	(Ameen et al., 2024; Berx et al., 2024; Blaurock et al., 2024; Calitz et al., 2017; Cheng et al., 2022; Ellwart et al., 2022; Frank et al., 2019; Hamza et al., 2024; Koponen et al., 2023; Le & Cayrat, 2024; Lehmann-Willenbrock & Hung, 2024; Lorson et al., 2023; Paluch et al., 2022; Schultz, 2021; Siemon, 2022; Terblanche et al., 2024; Vassilakopoulou et al., 2023; Zywiolek, 2024)
	Ethnography	0	

(*continued*)

Table 2. (*continued*)

Research Method	Research Design	Number of assigned studies	Assigned studies
	Case Studies	14	(Bauer et al., 2023; Ben-Assuli et al., 2023; Cabrero-Daniel et al., 2024; David et al., 2023; Fossum et al., 2018; Gidel et al., 2020; Kokotinis et al., 2024; Li et al., 2023; X. Liu et al., 2024; Loske & Klumpp, 2021b, 2021a; Malik et al., 2020; Peruzzini et al., 2024; Stadnicka & Antonelli, 2019)
Mixed Methods	Convergent Mixed Methods	0	
	Explanatory Sequential Mixed Methods	1	(Siemon, 2022)
	Exploratory Sequential Mixed Methods	6	(Ameen et al., 2024; Berx et al., 2024; Blaurock et al., 2024; Cheng et al., 2022; Ellwart et al., 2022; Zywiolek, 2024)

3.1 Challenges in Managing and Organizing Human-AI Collaboration

Integrating AI into organizations offers substantial potential for value creation but also presents numerous management and organizational challenges, impacting areas such as team collaboration, leadership, and organizational culture. This review identifies five main categories of these challenges:

- Hybrid human-AI teams (16 studies) - this includes designing team structures (A. Dennis et al., 2023; Hamza et al., 2024; Jain et al., 2023; Siemon, 2022), understanding collaboration dynamics (Ameen et al., 2024; Cabrero-Daniel et al., 2024; Chala & Graß, 2024; Kridalukmana et al., 2019; Lehmann-Willenbrock & Hung, 2024; Vassilakopoulou et al., 2023), fostering trust with AI teammates (Georganta & Ulfert, 2024a, 2024b; Gidel et al., 2020; Hoch et al., 2024), and managing delegation between humans and AI (Adam et al., 2024; Fügener et al., 2021). Studies underline the complexities of integrating AI into traditional teams, requiring careful design to balance efficiency, trust, and human autonomy. Key issues include clearly defining the roles of AI as team members and ensuring human oversight to avoid inefficiencies or over-reliance on AI.
- AI integration in work processes (14 studies) - this explores the broader organizational implications of incorporating AI, including employee outcomes (Ben-Assuli et al., 2023; Blaurock et al., 2024; Bouarfa et al., 2018; Franken et al., 2023; Gnewuch et al., 2023; Xu et al., 2024), ethical considerations (Gonzalez-Jimenez & Pinto, 2024; Haupt et al., 2024), and defining the roles of humans and AI within workflows

(Koponen et al., 2023; Loske & Klumpp, 2021b; Q. Wang et al., 2023) and organizational processes (Ben-Assuli et al., 2023; Kattel et al., 2020; Yu et al., 2021). The emergence of ethical challenges underscores the growing need for responsible AI implementation. Redesigning workflows and task allocation is necessary to optimize the distribution of tasks between AI and human workers, ensuring a balance that avoids increasing employee stress or workload.

- Human-AI decision-making (7 studies) - this underscores the growing reliance on AI for managerial tasks (Franken et al., 2023; Lehmann et al., 2022; Listopad, 2020; Loske & Klumpp, 2021a) and intelligent decision-making (Rodriguez & Hilaire, 2023; Rumovskaya, 2020; J. Wu et al., 2023). Maintaining trust requires balancing human oversight and reliance on AI to ensure that decisions are both informed and trustworthy.
- Trust in AI (11 studies) - this reflects ongoing concerns about fear of replacement by AI (Calitz et al., 2017; Ellwart et al., 2022; W. Wang et al., 2024; T.-J. Wu et al., 2024) and the need for transparency to mitigate biases (Cheng et al., 2022; Fügener et al., 2021; Pareek et al., 2024; Selmi et al., 2017; Terblanche et al., 2024; Zywiolek, 2024). Building trust in AI requires addressing negative emotions, enhancing the transparency of AI processes, and managing expectations about its capabilities. Organizations must implement responsible AI strategies to address job displacement and AI bias concerns.

Human-robot collaboration (HRC) (44 studies) - this emerges as the most extensively studied category, focusing on optimizing interactions in shared human-robot environments (Azadeh et al., 2023; Bauer et al., 2023a, 2023b; Ellwart et al., 2022; Fossum et al., 2018; Frank et al., 2019; Lin et al., 2022; Lin & Lukodono, 2022; X. Liu et al., 2024; Malik et al., 2020; Mao et al., 2024a; Oliff et al., 2020; Srinivas & Yu, 2022) and addressing practical and safety concerns in collaborative settings (Berx et al., 2024; Chu et al., 2024; Ferreira et al., 2021; Hu et al., 2020; Y. Liu et al., 2023; Pettersson & Falkman, 2024; Zanchettin & Lacevic, 2022; K. Zhang et al., 2020; Zhu et al., 2024). Task planning and allocation (Karuppiah et al., 2023; Kokotinis et al., 2024; Li et al., 2023; Lorson et al., 2023; Lv et al., 2021; Mao et al., 2024b; Nourmohammadi et al., 2022; Peng et al., 2024; Peruzzini et al., 2024; Stadnicka & Antonelli, 2019; Weckenborg et al., 2020; Yu et al., 2021; M. Zhang et al., 2022a, 2022b) and experience of humans collaborating with machines (Fan & Zheng, 2024; Le & Cayrat, 2024; Molitor & Renkema, 2022; Paluch et al., 2022; Pasparakis et al., 2023; Schultz, 2021; Shanks et al., 2024; T. Wu et al., 2022) were also very often raised in this category. In collaborative robotics environments, ensuring worker safety, synchronization, and risk mitigation are critical.

Further challenges arise in adapting organizational structures and workflows. These also include managing human factors and employee well-being, along with addressing ethical and cultural considerations. The literature review highlights the importance of clearly defining roles, balancing automation with human oversight, and addressing ethical concerns for the successful adoption of AI. Overcoming resistance to change requires effectively demonstrating the tangible benefits of AI for both employees and the organization.

3.2 Research Approaches in Human-AI Collaboration Studies

The review systematizes research methods employed in studies of human-AI collaboration, drawing on the theoretical framework by Creswell & Creswell (2018) and extending it to include simulation as a research design. The research methods are divided into quantitative, qualitative, and mixed methods, with quantitative methods being more frequently used than qualitative ones (Table 2).

In quantitative methods, survey research includes cross-sectional studies (7 studies) and longitudinal studies (1 study). Experimental research is the most frequently applied quantitative method, encompassing true experiments (19 studies), quasi-experiments (15 studies), and simulation (30 studies), with simulation also referred to as a computational experiment.

Within qualitative methods, grounded theory includes 18 studies, and case studies include 14. Data collection for qualitative methods typically involves observations, workshops, interviews, and open-ended surveys.

In mixed methods, the exploratory sequential design is more commonly adopted than the explanatory sequential design (6 studies). It starts with a grounded theory design, followed by either a cross-sectional study or a design from the experimental research group. The explanatory sequential design includes 1 study.

An analysis of the relative frequency of research designs in Human-Robot Collaboration and non-HRC studies indicates that simulation design is frequently used in HRC studies, accounting for nearly half of them. In contrast, non-HRC studies show a more balanced distribution of research designs, with true experiments being the most common, appearing in almost one-third of these studies. Survey research is rarely used in HRC studies but is present in over 10% of non-HRC studies. Qualitative methods are applied at similar rates in both HRC and non-HRC studies.

4 Discussion

This systematic literature review identified key management and organizational challenges in human-AI collaboration, along with the research methods employed in this field. These challenges were categorized into five key areas: Hybrid Human-AI Teams, AI Integration in Work Processes, Trust in AI, Human-AI Decision-Making, and Human-Robot Collaboration (HRC). These categories highlight the diverse and complex interactions between humans and AI in organizational settings.

Hybrid Human-AI Teams emerged as a central topic, particularly regarding role definition, task delegation, and trust-building. Consistent with previous studies (Seeber et al., 2020), this review emphasizes the importance of clearly defining human and AI roles to enhance efficiency and reduce miscommunication.

AI Integration in Work Processes presents challenges related to workflow adaptation and task allocation. While prior research (Dwivedi et al., 2021) has explored AI's potential to optimize operations, this review highlights the managerial complexity involved in balancing AI capabilities with human oversight, particularly in preventing increased workload and stress.

Trust in AI remains a fundamental concern in human-AI collaboration. Transparency and accountability are key to fostering trust and reducing resistance to AI adoption (Abedin, 2022). This review underscores the importance of addressing employee anxieties about AI's impact on job security and work environments.

Human-AI Decision-Making is a relatively underexplored area, pointing to a gap in research. While prior studies (Raisch & Krakowski, 2021) have emphasized the need for empirical investigations into AI-assisted decision-making, this review provides preliminary insights into how AI-generated recommendations influence human judgment and leadership dynamics.

Human-robot collaboration (HRC) emerged as the most extensively studied category, with a strong focus on workflow optimization, safety, and productivity (McGirr et al., n.d.). This study extends previous findings by examining HRC within broader organizational and managerial contexts, emphasizing the role of trust, ethics, and workforce well-being in addition to technical considerations.

Beyond HRC, the literature encompasses a wider range of topics, including team dynamics, task delegation, and AI-supported decision-making. Research Lane & Saint-Martin (2021) suggests that AI is more likely to reorganize tasks rather than replace jobs entirely, positioning AI as a tool that enhances human productivity rather than a substitute for human labor.

5 Conclusions

Contrary to initial expectations that HRC would be a more niche topic, this review reveals that human-AI collaboration research is currently dominated by HRC studies. Furthermore, the level of AI autonomy observed in existing research is generally lower than anticipated, with a prevalence of studies focusing on reactive AI tools rather than fully autonomous, goal- or utility-based agents. Future research is expected to shift toward higher autonomy, goal-driven AI, and more advanced software actuators, enabling deeper human-AI interactions.

This study contributes to the field by categorizing key management and organizational challenges in human-AI collaboration and identifying gaps in the literature. By examining hybrid teams, AI integration, trust, decision-making, and HRC, the review provides a structured framework for understanding human-AI interactions in organizational settings. Unlike previous systematic reviews that focus primarily on technical aspects, this study expands the discussion to managerial and organizational perspectives, incorporating trust, ethics, and cultural considerations. Additionally, the analysis of hybrid teams offers new insights into role definition and team dynamics, which remain underexplored in prior research.

References

Abedin, B.: Managing the tension between opposing effects of explainability of artificial intelligence: a contingency theory perspective. Internet Res. **32**(2), 425–453 (2022). https://doi.org/10.1108/INTR-05-2020-0300

Adam, M., Diebel, C., Goutier, M., Benlian, A.: Navigating autonomy and control in human-AI delegation: user responses to technology- versus user-invoked task allocation. Decision Support Syst. **180**. https://doi.org/10.1016/j.dss.2024.114193

Ameen, N., Pagani, M., Pantano, E., Cheah, J., Tarba, S., Xia, S.: The rise of human-machine collaboration: managers' perceptions of leveraging artificial intelligence for enhanced B2B service recovery. Br. J. Manag. (2024). https://doi.org/10.1111/1467-8551.12829

Azadeh, K., Roy, D., de Koster, R., Khalilabadi, S.: Zoning strategies for human-robot collaborative picking. Decis. Sci. (2023). https://doi.org/10.1111/deci.12620

Bauer, J.C., Abdous, M.-A., Kurscheid, S., Lucas, F., Lozenguez, G., Daub, R.: Towards Enabling Human-Robot Collaboration In Industry: Identification Of Current Implementation Barriers. In Herberger D. & Hubner M. (Eds.), *Proc. Conf. Prod. Syst. Logist.,* pp. 67–78. Publishing in cooperation with TIB - Leibniz Information Centre for Science and Technology University Library; Scopus (2023). https://doi.org/10.15488/15254

Ben-Assuli, O., Heart, T., Klempfner, R., Padman, R.: Human-machine collaboration for feature selection and integration to improve congestive Heart failure risk prediction. Decision Support Syst. **172** (2023). https://doi.org/10.1016/j.dss.2023.113982

Berx, N., Decré, W., Pintelon, L.: A tool to evaluate industrial cobot safety readiness from a system-wide perspective: an empirical validation. Safety Sci. **170** (2024). https://doi.org/10.1016/j.ssci.2023.106380

Bilan, S., Šuleř, P., Skrynnyk, O., Krajňáková, E., Vasilyeva, T.: Systematic bibliometric review of artificial intelligence technology in organizational management, development, change and culture. Business: Theory Practice **23**(1), 1–13 (2022). https://doi.org/10.3846/BTP.2022.13204

Blaurock, M., Büttgen, M., & Schepers, J.: Designing collaborative intelligence systems for employee-AI service co-production. J. Serv. Res. Scopus (2024). https://doi.org/10.1177/10946705241238751

Bouarfa, S., Müller, J., Blom, H.: Evaluation of a Multi-Agent System approach to airline disruption management. J. Air Transport Manage. **71**, 108–118. Scopus (2018). https://doi.org/10.1016/j.jairtraman.2018.05.009

Cabrero-Daniel, B., Herda, T., Pichler, V., Eder, M.: Exploring human-AI collaboration in agile: customised LLM meeting assistants. In: Šmite D., Guerra E., Wang X., Marchesi M., Gregory P. (eds.) Lect. Notes Bus. Inf. Process., vol. 512 LNBIP, pp. 163–178. Springer Science and Business Media Deutschland GmbH; Scopus (2024). https://doi.org/10.1007/978-3-031-61154-4_11

Calitz, A., Poisat, P., Cullen, M.: The future African workplace: The use of collaborative robots in manufacturing. SA J. Human Resource Manage. **15** (2017). https://doi.org/10.4102/sajhrm.v15i0.901

Chala, S.A., Graß, A.: Human-AI interaction for semantic knowledge enrichment of AI model output. In: Artif. Intell. In Manuf.: Enabling Intell., Flex. And Cost-Eff. Prod. Through AI, pp. 43–54. Springer Nature; Scopus (2024). https://doi.org/10.1007/978-3-031-46452-2_3

Chand, S., McDaid, A., Lu, Y.: Dynamic muscle fatigue assessment using s-EMG technology towards human-centric human-robot collaboration. J. Manuf. Syst. **68**, 508–522 (2023). https://doi.org/10.1016/j.jmsy.2023.05.022

Cheng, X., Zhang, X., Yang, B., Fu, Y.: An investigation on trust in AI-enabled collaboration: Application of AI-Driven chatbot in accommodation-based sharing economy. Electronic Commerce Res. Appl. **54**. Scopus (2022). https://doi.org/10.1016/j.elerap.2022.101164

Chu, H., Zhang, T., Zou, Y., Sun, H.: "Follower" to "Collaborator": a robot proactive collaborative controller based on human multimodal information for 3D handling/ assembly scenarios. J. Manuf. Syst. **73**, 287–306 (2024). https://doi.org/10.1016/j.jmsy.2024.02.008

Cimpeanu, T., Han, T., IEEE.: Making an Example: Signalling Threat in the Evolution of Cooperation. University of Teesside. In: 2020 IEEE Congress on Evolutionary Computation (CEC) (2020)

Creswell, J.W., Creswell, J.D.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2018)

David, J., Coatanéa, E., Lobov, A.: Deploying OWL ontologies for semantic mediation of mixed-reality interactions for human-robot collaborative assembly. J. Manuf. Syst. **70**, 359–381 (2023). https://doi.org/10.1016/j.jmsy.2023.07.013

Davis, J.P., Eisenhardt, K.M., Bingham, C.B.: Developing theory through simulation methods. Acad. Manag. Rev. **32**(2), 480–499 (2007). https://doi.org/10.5465/amr.2007.24351453

Dennis, A., Lakhiwal, A., Sachdeva, A.: AI agents as team members: effects on satisfaction, conflict, trustworthiness, and willingness to work with. J. Manag. Inf. Syst. **40**(2), 307–337 (2023). https://doi.org/10.1080/07421222.2023.2196773

Dwivedi, Y.K., et al.: Artificial Intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. Int. J. Inf. Manage. **57**, 101994 (2021). https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Ellwart, T., Schauffel, N., Antoni, C. H., & Timm, I. J. (2022). I vs. Robot: Sociodigital self-comparisons in hybrid teams from a theoretical, empirical, and practical perspective. Gruppe. Interaktion. Organisation. Zeitschrift Fur Angewandte Organisationspsychologie **53**(3), 273–284. Scopus. https://doi.org/10.1007/s11612-022-00638-5

Fan, J., Zheng, P.: A vision-language-guided robotic action planning approach for ambiguity mitigation in human-robot collaborative manufacturing. J. Manuf. Syst. **74**, 1009–1018 (2024). https://doi.org/10.1016/j.jmsy.2024.05.003

Ferreira, C., Figueira, G., Amorim, P.: Scheduling Human-Robot Teams in collaborative working cells. Int. J. Prod. Econ. **235** (2021). https://doi.org/10.1016/j.ijpe.2021.108094

Fossum, K., Danielsen, B., Aarseth, W., Johnsen, S.: A project management issue of new technology developments: a case study on lack of human factors' attention in human-robot interaction. Proceedings of the Institution of Mechanical Engineers Part O-Journal of Risk and Reliability **232**(2), 164–173 (2018). https://doi.org/10.1177/1748006X17728601

Frank, M., Ruvald, R., Johansson, C., Larsson, T., Larsson, A.: Towards autonomous construction equipment-supporting on-site collaboration between automatons and humans. Int. J. Product Dev. **23**(4), 292–308 (2019). Scopus. https://doi.org/10.1504/IJPD.2019.105496

Franken, A., Manthei, C., Senderek, R., Stich, V.: Demands and challenges for SME regarding the human-centered implementation of innovative technologies and AI. In: Herberger, D., Hübner, M., Stich, V. (eds.) Proc. Conf. Prod. Syst. Logist., pp. 718–727. Publishing in cooperation with TIB - Leibniz Information Centre for Science and Technology University Library; Scopus. https://doi.org/10.15488/13491

Fügener, A., Grahl, J., Gupta, A., Ketter, W.: WILL humans-in-the-loop become borgs? merits and pitfalls of working with AI. Mis Quarterly **45**(3), 1527–1556 (2021). https://doi.org/10.25300/MISQ/2021/16553

Fügener, A., Grahl, J., Gupta, A., Ketter, W.: Cognitive challenges in human-artificial intelligence collaboration: investigating the path toward productive delegation. Inf. Syst. Res. **33**(2), 678–696 (2022). https://doi.org/10.1287/isre.2021.1079

Gaczek, P., Leszczynski, G., Mouakher, A.: Collaboration with machines in B2B marketing: overcoming managers' aversion to AI-CRM with explainability. Ind. Mark. Manage. **115**, 127–142 (2023). https://doi.org/10.1016/j.indmarman.2023.09.007

Gebru, B., Zeleke, L., Blankson, D., Nabil, M., Nateghi, S., Homaifar, A., Tunstel, E.: A review on human–machine trust evaluation: human-centric and machine-centric perspectives. IEEE Trans. Hum.-Mach. Syst. **52**(5), 952–962 (2022). IEEE Transactions on Human-Machine Systems. https://doi.org/10.1109/THMS.2022.3144956

Georganta, E., Ulfert, A.: My colleague is an AI! Trust differences between AI and human teammates. Team Performance Manage. **30**(1/2), 23–37 (2024). https://doi.org/10.1108/TPM-07-2023-0053

Georganta, E., Ulfert, A.: Would you trust an AI team member? team trust in human-AI teams. J. Occupational Organizational Psychol. **97**(3), 1212–1241 (2024). https://doi.org/10.1111/joop.12504

Gidel, T., et al.: Interaction model and respect of rules to enhance collaborative brainstorming results. Advances in Science, Technology and Engineering Systems, 5(2), 484–493 (2020). Scopus. https://doi.org/10.25046/aj050262

Gnewuch, U., Morana, S., Hinz, O., Kellner, R., Maedche, A.: More Than a bot? the impact of disclosing human involvement on customer interactions with hybrid service agents. Information Systems Research (2023). https://doi.org/10.1287/isre.2022.0152

Gonzalez-Jimenez, H., Pinto, D.: Can AI robots foster social inclusion? Exploring the role of immersive augmentation in hospitality. Int. J. Contemporary Hospitality Manage. (2024). https://doi.org/10.1108/IJCHM-09-2023-1459

Hamza, M., Siemon, D., Akbar, M.A., Rahman, T.: Human-AI collaboration in software engineering: lessons learned from a hands-on workshop. In: Proc. - IEEE/ACM Int. Workshop Softw.-Intensive Bus., IWSiB, 7–14 (2024). Scopus. https://doi.org/10.1145/3643690.3648236

Happach, R.M., Tilebein, M.: Simulation as research method: modeling social interactions in management science. In: Misselhorn, C. (ed.) Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation, pp. 239–259. Springer (2015). https://doi.org/10.1007/978-3-319-15515-9_13

Haupt, M., Freidank, J., Haas, A.: Consumer responses to human-AI collaboration at organizational frontlines: strategies to escape algorithm aversion in content creation. RMS (2024). https://doi.org/10.1007/s11846-024-00748-y

Hoch, T., et al.: Multi-stakeholder perspective on human-AI collaboration in industry 5.0. In: Artif. Intell. In Manuf.: Enabling Intell., Flex. And Cost-Eff. Prod. Through AI, pp. 407–421. Springer Nature; Scopus. https://doi.org/10.1007/978-3-031-46452-2_23

Hu, D., Li, S., Cai, J., Hu, Y., IEEE.: Toward intelligent workplace: prediction-enabled proactive planning for human-robot coexistence on unstructured construction sites. University of Tennessee System, 2412–2423 (2020). https://doi.org/10.1109/WSC48552.2020.9384077

Jain, R., Garg, N., Khera, S.N.: Comparing differences of trust, collaboration and communication between human-human vs human-bot teams: an experimental study. CERN IdeaSquare J. Exp. Innov. **7**(2), 8–16 (2023). Scopus. https://doi.org/10.23726/cij.2022.1387

Karuppiah, K., Sankaranarayanan, B., Ali, S., Bhalaji, R.: Decision modeling of the challenges to human-robot collaboration in industrial environment: A real world example of an emerging economy. Flexible Serv. Manuf. J. **35**(4), 1007–1037 (2023). https://doi.org/10.1007/s10696-022-09474-7

Kattel, R., Lember, V., Tonurist, P.: Collaborative innovation and human-machine networks. Public Manag. Rev. **22**(11), 1652–1673 (2020). https://doi.org/10.1080/14719037.2019.1645873

Klatt, D.F.: Business Systematic Literature Reviews (2023)

Kokotinis, G., Michalos, G., Arkouli, Z., Makris, S.: A behavior trees-based architecture towards operation planning in hybrid manufacturing. Int. J. Comput. Integr. Manuf. **37**(3), 324–349 (2024). https://doi.org/10.1080/0951192X.2023.2228254

Koponen, J., Julkunen, S., Laajalahti, A., Turunen, M., Spitzberg, B.: Work Characteristics Needed by Middle Managers When Leading AI-Integrated Service Teams. Journal of Service Research. Scopus (2023). https://doi.org/10.1177/10946705231220462

Kridalukmana, R., Lu, H., Naderpour, M.: Component-Based Transparency to Comprehend Intelligent Agent Behaviour for Human-Autonomy Teaming. In: Zou L., Fang L., Fu B., Niu P. (eds.) Proc. IEEE Int. Conf. Intell. Syst. Knowl. Eng., *ISKE,* pp. 771–777. Institute of Electrical

and Electronics Engineers Inc.; Scopus (2019). https://doi.org/10.1109/ISKE47853.2019.9170206

Łabędzki, R.: Hybrid Multi-Agent Systems as a subject of scientific research in the field of management. (2024). https://doi.org/10.2139/ssrn.4985522

Lane, M., Saint-Martin, A.: The impact of Artificial Intelligence on the labour market: What do we know so far? (OECD Social, Employment and Migration Working Papers 256; OECD Social, Employment and Migration Working Papers, vol. 256 (2021). https://doi.org/10.1787/7c895724-en

Le, K., Cayrat, C.: Howdy, Robo-Partner: Exploring artificial companionship and its stress-alleviating potential for service employees. J. Serv. Manage. (2024). https://doi.org/10.1108/JOSM-09-2023-0408

Lehmann, C., Haubitz, C., Fügener, A., Thonemann, U.: The risk of algorithm transparency: how algorithm complexity drives the effects on the use of advice. Production Oper. Manage. **31**(9), 3419–3434 (2022). https://doi.org/10.1111/poms.13770

Lehmann-Willenbrock, N., Hung, H.: A multimodal social signal processing approach to team interactions. Organ. Res. Methods **27**(3), 477–515 (2024). https://doi.org/10.1177/10944281231202741

Li, S., Zheng, P., Pang, S., Wang, X., Wang, L.: Self-organising multiple human-robot collaboration: a temporal subgraph reasoning-based method. J. Manuf. Syst. **68**, 304–312 (2023). https://doi.org/10.1016/j.jmsy.2023.03.013

Lin, C., Lukodono, R.: Classification of mental workload in Human-robot collaboration using machine learning based on physiological feedback. J. Manuf. Syst. **65**, 673–685 (2022). https://doi.org/10.1016/j.jmsy.2022.10.017

Lin, C., Wang, K., Tadesse, A., Woldegiorgis, B.: Human-robot collaboration empowered by hidden semi-Markov model for operator behaviour prediction in a smart assembly system. J. Manuf. Syst. **62**, 317–333 (2022). https://doi.org/10.1016/j.jmsy.2021.12.001

Listopad, S.: Solving the regional power grid restoration problem with the prototype of the hybrid intelligent multi-agent system of heterogeneous thinking. In: Proc. - Int. Conf. Control Syst., Math. Model., Autom. Energy Effic., SUMMA, pp. 243–247. Scopus (2020). https://doi.org/10.1109/SUMMA50634.2020.9280796

Liu, T., et al.: Toward fast 3D human activity recognition: a refined feature based on minimum joint freedom model (Mint). J. Manuf. Syst. **66**, 127–141 (2023). https://doi.org/10.1016/j.jmsy.2022.11.009

Liu, X., Li, G., Xiang, F., Tao, B., Jiang, G.: Blockchain-based cloud-edge collaborative data management for human-robot collaboration digital twin system. J. Manuf. Syst. **77**, 228–245 (2024). https://doi.org/10.1016/j.jmsy.2024.09.006

Liu, Y., Liu, K., Wang, G., Sun, Z., Jin, L.: Noise-tolerant zeroing neurodynamic algorithm for upper limb motion intention-based human-robot interaction control in non-ideal conditions. Expert Syst. Appl. **213** (2023). https://doi.org/10.1016/j.eswa.2022.118891

Lorson, F., Fügener, A., Hübner, A.: New team mates in the warehouse: Human interactions with automated and robotized systems. IISE Trans. **55**(5), 536–553 (2023). https://doi.org/10.1080/24725854.2022.2072545

Loske, D., Klumpp, M.: Human-AI collaboration in route planning: an empirical efficiency-based analysis in retail logistics. Int. J. Prod. Econ. **241** (2021a). https://doi.org/10.1016/j.ijpe.2021.108236

Loske, D., Klumpp, M.: Intelligent and efficient? an empirical analysis of human-AI collaboration for truck drivers in retail logistics. Int. J. Logist. Manag. **32**(4), 1356–1383 (2021). https://doi.org/10.1108/IJLM-03-2020-0149

Lv, Q., Zhang, R., Sun, X., Lu, Y., Bao, J.: A digital twin-driven human-robot collaborative assembly approach in the wake of COVID-19. J. Manuf. Syst. **60**, 837–851 (2021). https://doi.org/10.1016/j.jmsy.2021.02.011

Madanchian, M., Taherdoost, H., Mohamed, N.: AI-based human resource management tools and techniques; a systematic literature review. Procedia Comput. Sci. **229**, 367–377 (2023). https://doi.org/10.1016/j.procs.2023.12.039

Malik, A., Masood, T., Bilberg, A.: Virtual reality in manufacturing: Immersive and collaborative artificial-reality in design of human-robot workspace. Int. J. Comput. Integr. Manuf. **33**(1), 22–37 (2020). https://doi.org/10.1080/0951192X.2019.1690685

Mao, Z., Sun, Y., Fang, K., Huang, D., Zhang, J.: Balancing and scheduling of assembly line with multi-type collaborative robots. Int. J. Prod. Econ. **271** (2024a). https://doi.org/10.1016/j.ijpe.2024.109207

Mao, Z., Sun, Y., Fang, K., Huang, D., Zhang, J.: Model and metaheuristic for human-robot collaboration assembly line worker assignment and balancing problem. Comput. Oper. Res. **165** (2024b). https://doi.org/10.1016/j.cor.2024.106605

Mao, Z., Zhang, J., Fang, K., Huang, D., Sun, Y.: Balancing U-type assembly lines with human-robot collaboration. Comput. Oper. Res. **159** (2023). https://doi.org/10.1016/j.cor.2023.106359

McGirr, L., Jin, Y., Price, M., West, A., van Lopik, K., McKenna, V.: Human robot collaboration: taxonomy of interaction levels in manufacturing. In: 54th International Symposium on Robotics (ISR Europe 2022): Proceedings, pp. 197–204. Institute of Electrical and Electronics Engineers Inc. (2022). https://ieeexplore.ieee.org/document/9861817

Mo, F., et al.: PLC orchestration automation to enhance human-machine integration in adaptive manufacturing systems. J. Manuf. Syst. **71**, 172–187 (2023). https://doi.org/10.1016/j.jmsy.2023.07.015

Molitor, M., Renkema, M.: Human-robot collaboration in a smart industry context: does HRM matter? In Adv. Ser. Manage. **28**, 105–123. Emerald Group Holdings Ltd.; Scopus (2022). https://doi.org/10.1108/S1877-636120220000028008

Nourmohammadi, A., Fathi, M., Ng, A.: Balancing and scheduling assembly lines with human-robot collaboration tasks. Comput. Operat. Res. **140** (2022). https://doi.org/10.1016/j.cor.2021.105674

Oliff, H., Liu, Y., Kumar, M., Williams, M., Ryan, M.: Reinforcement learning for facilitating human-robot-interaction in manufacturing. J. Manuf. Syst. **56**, 326–340 (2020). https://doi.org/10.1016/j.jmsy.2020.06.018

Page, M.J., et al.: The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, n71 (2021). https://doi.org/10.1136/bmj.n71

Paluch, S., Tuzovic, S., Holz, H., Kies, A., Jörling, M.: "My colleague is a robot"—Exploring frontline employees' willingness to work with collaborative service robots. J. Serv. Manag. **33**(2), 363–388 (2022). https://doi.org/10.1108/JOSM-11-2020-0406

Pareek, S., Velloso, E., Goncalves, J.: Trust development and repair in AI-assisted decision-making during complementary expertise. ACM Conf. Fairness, Account., Transpar., FAccT, pp. 546–561. Scopus (2024). https://doi.org/10.1145/3630106.3658924

Pasparakis, A., De Vries, J., De Koster, R.: Assessing the impact of human-robot collaborative order picking systems on warehouse workers. Int. J. Prod. Res. **61**(22), 7776–7790 (2023). https://doi.org/10.1080/00207543.2023.2183343

Peng, J., Kimmig, A., Wang, D., Niu, Z., Tao, X., Ovtcharova, J.: Intention recognition-based human-machine interaction for mixed flow assembly. J. Manuf. Syst. **72**, 229–244 (2024). https://doi.org/10.1016/j.jmsy.2023.11.021

Peruzzini, M., Prati, E., Pelicciari, M.: A framework to design smart manufacturing systems for Industry 5.0 based on the human-automation symbiosis. Int. J. Comput. Integrated Manuf. **37**(10–11), 1426–1443 (2024). https://doi.org/10.1080/0951192X.2023.2257634

Pettersson, J., Falkman, P.: Intended human arm movement direction prediction using eye tracking. Int. J. Comput. Integr. Manuf. **37**(9), 1107–1125 (2024). https://doi.org/10.1080/0951192X.2023.2229288

Prikshat, V., Islam, M., Patel, P., Malik, A., Budhwar, P., Gupta, S.: AI-Augmented HRM: Literature review and a proposed multilevel framework for future research. Technological Forecasting and Social Change, 193 (2023). https://doi.org/10.1016/j.techfore.2023.122645

Qamar, Y., Agrawal, R.K., Samad, T.A., Chiappetta Jabbour, C.J.: When technology meets people: the interplay of artificial intelligence and human resource management. J. Enterp. Inf. Manag. **34**(5), 1339–1370 (2021). https://doi.org/10.1108/JEIM-11-2020-0436

Raisch, S., Krakowski, S.: Artificial intelligence and management: the automation–augmentation paradox. Acad. Manag. Rev. **46**(1), 192–210 (2021). https://doi.org/10.5465/amr.2018.0072

Rodriguez, S., Hilaire, V.: A methodological approach for the analysis and design of Human-Swarm interactions based upon feedback loops. Expert Syst. Appl. **217** (2023). https://doi.org/10.1016/j.eswa.2022.119482

Rumovskaya, S.: Visualization of team cohesion in hybrid intelligent multi-agent systems. In: Proc. - Int. Conf. Control Syst., Math. Model., Autom. Energy Effic., SUMMA, pp. 620–623. Scopus (2020). https://doi.org/10.1109/SUMMA50634.2020.9280790

Schultz, C.: The future and the role of human resource management in South Africa during the Fourth Industrial Revolution. Sa J. Hum. Resource Manage. **19** (2021). https://doi.org/10.4102/sajhrm.v19i0.1624

Seeber, I., et al.: Machines as teammates: a research agenda on AI in team collaboration. Inf. Manage. **57**(2), 103174 (2020). https://doi.org/10.1016/j.im.2019.103174

Selmi, A., Brahmi, Z., Gammoudi, M.M.: PACT: a new trust prediction method for multi-agents recommender systems. In: Cellary, W., Fugini, M., Reddy S. (eds.) Proc. - IEEE Int. Conf. Enabling Technol.: Infrastruct. Collab. Enterp., WETICE, pp. 9–14). Institute of Electrical and Electronics Engineers Inc. Scopus. https://doi.org/10.1109/WETICE.2017.29

Shanks, I., Scott, M. L., Mende, M., van Doorn, J., Grewal, D.: Cobotic service teams and power dynamics: Understanding and mitigating unintended consequences of human-robot collaboration in healthcare services. Journal of the Academy of Marketing Science. Scopus (2024). https://doi.org/10.1007/s11747-024-01004-1

Siemon, D.: Elaborating team roles for artificial intelligence-based teammates in human-AI collaboration. Group Decision Negotiation **31**(5), 871–912 (2022). https://doi.org/10.1007/s10726-022-09792-z

Srinivas, S., Yu, S.: Collaborative order picking with multiple pickers and robots: Integrated approach for order batching, sequencing and picker-robot routing. Int. J. Prod. Econ. **254** (2022). https://doi.org/10.1016/j.ijpe.2022.108634

Stadnicka, D., Antonelli, D.: Human-robot collaborative work cell implementation through lean thinking. Int. J. Comput. Integr. Manuf. **32**(6), 580–595 (2019). https://doi.org/10.1080/0951192X.2019.1599437

Sutton, A., Clowes, M., Preston, L., Booth, A.: Meeting the review family: Exploring review types and associated information retrieval requirements. Health Info. Libr. J. **36**(3), 202–222 (2019). https://doi.org/10.1111/hir.12276

Taboada, I., Daneshpajouh, A., Toledo, N., de Vass, T.: Artificial intelligence enabled project management: a systematic literature review. Appl. Sci. **13**(8), 5014 (2023). https://doi.org/10.3390/app13085014

Terblanche, N.H.D., van Heerden, M., Hunt, R.: The influence of an artificial intelligence chatbot coach assistant on the human coach-client working alliance. Coaching **17**(2), 189–206 (2024). Scopus. https://doi.org/10.1080/17521882.2024.2304792

Vassilakopoulou, P., Haug, A., Salvesen, L., Pappas, I.: Developing human/AI interactions for chat-based customer services: lessons learned from the Norwegian government. Eur. J. Inf. Syst. **32**(1), 10–22 (2023). https://doi.org/10.1080/0960085X.2022.2096490

Vaswani, A., et al.: Attention is All you Need. Advances in Neural Information Processing Systems, 30 (2017). https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Votto, A.M., Valecha, R., Najafirad, P., Rao, H.R.: Artificial intelligence in tactical human resource management: a systematic literature review. Int. J. Inf. Manage. Data Insights **1**(2), 100047 (2021). https://doi.org/10.1016/j.jjimei.2021.100047

Wang, Q., Wan, Y., Feng, F.: Human-machine collaborative scoring of subjective assignments based on sequential three-way decisions. Expert Syst. Appl. **216** (2023). https://doi.org/10.1016/j.eswa.2022.119466

Wang, W., Gao, G., Agarwal, R.: Friend or foe? teaming between artificial intelligence and workers with variation in experience. Manage. Sci. **70**(9), 5753–5775 (2024). https://doi.org/10.1287/mnsc.2021.00588

Weckenborg, C., Kieckhäfer, K., Müller, C., Grunewald, M., & Spengler, T.S.: Balancing of assembly lines with collaborative robots. Business Res. **13**(1), 93–132 (2020). Scopus. https://doi.org/10.1007/s40685-019-0101-y

Wu, J., Wu, Y., Chen, K., Hua, L.: Building socially intelligent AI systems: evidence from the trust game using artificial agents with deep learning. Manage. Sci. **69**(12), 7236–7252 (2023). https://doi.org/10.1287/mnsc.2023.4782

Wu, T., Li, J., Wu, Y.: Employees' job insecurity perception and unsafe behaviours in human-machine collaboration. Manag. Decis. **60**(9), 2409–2432 (2022). https://doi.org/10.1108/MD-09-2021-1257

Wu, T.-J., Liang, Y., & Wang, Y. (2024). The Buffering Role of Workplace Mindfulness: How Job Insecurity of Human-Artificial Intelligence Collaboration Impacts Employees' Work–Life-Related Outcomes. *Journal of Business and Psychology*. Scopus. https://doi.org/10.1007/s10869-024-09963-6

Xia, X., Guo, Y., Wang, Y., Yang, Y., Shi, Y., Men, H.: Advancing cross-subject olfactory EEG recognition: a novel framework for collaborative multimodal learning between human-machine. Expert Syst. Appl. **250** (2024). https://doi.org/10.1016/j.eswa.2024.123972

Xu, J., Tang, X., Chang, E.-C., & Peng, H. (2024). Working with AI: the impact of organizational intelligent service strategy on employees' perception of career achievement. *Humanities and Social Sciences Communications*, *11*(1). Scopus. https://doi.org/10.1057/s41599-024-03265-1

Yu, T., Huang, J., Chang, Q.: Optimizing task scheduling in human-robot collaboration with deep multi-agent reinforcement learning. J. Manuf. Syst. **60**, 487–499 (2021). https://doi.org/10.1016/j.jmsy.2021.07.015

Yun, H., Jun, M.: Immersive and interactive cyber-physical system (I2CPS) and virtual reality interface for human involved robotic manufacturing. J. Manuf. Syst. **62**, 234–248 (2022). https://doi.org/10.1016/j.jmsy.2021.11.018

Zanchettin, A., Lacevic, B.: Safe and minimum-time path-following problem for collaborative industrial robots. J. Manuf. Syst. **65**, 686–693 (2022). https://doi.org/10.1016/j.jmsy.2022.10.020

Zhang, G., Chong, L., Kotovsky, K., Cagan, J.: Trust in an AI versus a Human teammate: The effects of teammate identity and performance on Human-AI cooperation. Computers in Human Behavior, 139. Scopus (2023). https://doi.org/10.1016/j.chb.2022.107536

Zhang, K., Xu, W., Yao, B., Ji, Z., Hu, Y., Feng, H., & IEEE.: Human Motion Recognition for Industrial Human-Robot Collaboration based on a Novel Skeleton Descriptor. Wuhan University of Technology, 404–410 (2020). https://doi.org/10.1109/case48305.2020.9216971

Zhang, M., Li, C., Shang, Y., Huang, H., Zhu, W., Liu, Y.: A task scheduling model integrating micro-breaks for optimisation of job-cycle time in human-robot collaborative assembly cells. Int. J. Prod. Res. **60**(15), 4766–4777 (2022). https://doi.org/10.1080/00207543.2021.1937746

Zhang, R., Lv, J., Li, J., Bao, J., Zheng, P., Peng, T.: A graph-based reinforcement learning-enabled approach for adaptive human-robot collaborative assembly operations. J. Manuf. Syst. **63**, 491–503 (2022). https://doi.org/10.1016/j.jmsy.2022.05.006

Zhang, T., Chu, H., Zou, Y., Sun, H.: A robust electromyography signals-based interaction interface for human-robot collaboration in 3D operation scenarios. Expert Syst. Appl. **238** (2024). https://doi.org/10.1016/j.eswa.2023.122003

Zhu, C., Yu, T., Chang, Q.: Task-oriented safety field for robot control in human-robot collaborative assembly based on residual learning. Expert Syst. Appl. **238** (2024). https://doi.org/10.1016/j.eswa.2023.121946

Zywiolek, J.: Building Trust in AI-Human Partnerships: Exploring Preferences and Influences in the Manufacturing Industry. Manage. Syst. Prod. Eng. **32**(2), 244–251 (2024). Scopus. https://doi.org/10.2478/mspe-2024-0024

Public Perceptions of Generative AI: Insights from Social Q&A Platforms

Jiaqi Liao[1](✉) and Chei Sian Lee[2]

[1] School of Information Management, Sun Yat-Sen University, Guangzhou 510000, China
liaojq7@mail2.sysu.edu.cn

[2] Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718, Singapore
LeeCS@ntu.edu.sg

Abstract. Generative AI (GenAI) discussions on social Q&A platforms reveal diverse public perceptions and user engagement shaped by platform-specific characteristics. This study analyzed discussions on Zhihu, Quora, and Reddit using topic modeling, topic similarity analysis and sentiment analysis, uncovering distinct user priorities and emotional responses. Zhihu focused on content creation, ethics, and personalized learning, with polarized sentiments. Quora emphasized business strategies, customer support, and technical learning, showing predominantly positive emotions. Reddit explored a wide range of topics, including communication tools and software development, with balanced sentiments. Cross-platform analysis revealed shared interests in communication, translation, and education, with notable overlaps in personalized learning and technical applications. These insights highlight how user demographics influence GenAI discourse, offering practical guidance for tailoring tools to regional needs. The comparative results have both practical and theoretical implications, offering valuable insights for stakeholders to tailor applications and strategies for different user groups related to GenAI.

Keywords: Generative artificial intelligence · Public discourse · Social Q&A · Topic modeling · Sentiment analysis · Cross-platform comparison

1 Background

Discussions on generative AI (GenAI) across social Q&A platforms provide valuable insights into public perceptions and engagement [1]. However, few studies have systematically explored these discourses. While GenAI applications in software development, business, and creative industries are expanding, concerns about content trustworthiness and societal impact persist [2]. Social Q&A platforms like Zhihu, Quora, and Reddit foster diverse, knowledge-driven discussions, shaped by linguistic and cultural contexts [3]. This study employs topic modeling and sentiment analysis to examine how users across these platforms discuss and perceive GenAI, identifying key thematic differences and shared concerns.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 349–358, 2025.
https://doi.org/10.1007/978-3-031-94171-9_33

2 Methods

2.1 Data Collection

GenAI-related data were collected from three popular social Q&A platforms, Zhihu, Quora, and Reddit, representing distinct online communities, each representing distinct online communities with varying linguistic, cultural, and user engagement characteristics [3, 4]. Data collection was conducted using predefined keywords, including "Generative artificial intelligence," "GPT," and "AI tools." The dataset included 11,320 responses from Zhihu, 2,225 from Quora, and 4,893 from Reddit. Topic modeling identified six themes on Zhihu and five each on Quora and Reddit, revealing overlaps and platform-specific interests (Table 1). Cross-platform topic similarities were analyzed using word embeddings and cosine similarity scores, while kernel density plots visualized sentiment distributions across platforms.

2.2 Data Analysis

To analyze the collected discussions, this study employed three key analytical techniques. First, topic modeling using Latent Dirichlet Allocation (LDA) was applied to identify major discussion themes across platforms. Second, topic similarity analysis leveraged word embeddings and cosine similarity scores to measure thematic overlaps between platforms. Finally, sentiment analysis was conducted using machine learning-based sentiment classification models, including the Baidu NLP API for Chinese text and Google Cloud NLP API for English text to evaluate sentiment trends and distribution across platforms.

3 Results

3.1 Topic Modeling

The LDA model was used from sklearn to extract topics from Zhihu, Quora, and Reddit, determining the optimal number of topics using perplexity scores. LDA is a generative probabilistic model used to identify topics within a collection of documents by finding groups of words that frequently occur together [5]. The lower scores indicate better generalization [6]. The results identified six topics for Zhihu, and five each for Quora and Reddit (Table 1).

Zhihu discussions focused on content creation (32.4%), ethical and societal issues (28.9%), communication (13.8%), learning and education (11.6%), research and software development (7.5%), and marketing and business (5.8%). Quora emphasized marketing and business (28.3%), customer support (27%), communication (25.4%), content creation (10.7%), and technical learning (8.6%). Reddit discussions centered on communication (22.7%), learning and education for translation (20.8%), software development (20.7%), personalized learning (19.8%), and marketing & business for human resource (16.0%).

Table 1. Topic distribution in three platforms.

Platform	Topic	Word stems associated with topics	Responses Count (Percentage)
Zhihu (n = 11,320)	Topic 0: Content Creation	Code, tools, content, images, video, work, plugins, search, translation, article	3669 (32.4%)
	Topic 1: Ethical and Societal	Human, development, work, world, society, replace, machine, search engine, logic, technology	3275 (28.9%)
	Topic 2: Communication	User, language, learning, dialogue, voice, natural language, chat, audio, interaction, real-time	1563 (13.8%)
	Topic 3: Learning and Education	Thesis, research, learning, article, work, story, content, writing, teacher, student	1310 (11.6%)
	Topic 4: Research and Software Development	Data, training, task, input, output, code, algorithm, calculation, testing, system	851 (7.5%)
	Topic 5: Marketing and Business	Product, company, industry, market, enterprise, customer, e-commerce, platform, investment, brand	652 (5.8%)
Quora (n = 2,225)	Topic 0: Marketing and Business	Marketing, blog, writing, post, media, social, audience, article, engage, topic	630 (28.3%)
	Topic 1: Customer Support	Service, business, chatbot, customer, online, platform, virtual, client, support, sell	601 (27%)

(*continued*)

Table 1. (*continued*)

Platform	Topic	Word stems associated with topics	Responses Count (Percentage)
	Topic 2: Communication	Language, text, chat, natural, humanlike, conversation, input, transformer, translation, chatbot	565 (25.4%)
	Topic 3: Content Creation	Creative, intelligence, produce, text, images, music, ideas, design, art, patterns	237 (10.7%)
	Topic 4: Learning and Education (Technical)	Search, learn, program, course, knowledge, computer, workshop, video, code, read	192 (8.6%)
Reddit (n = 4,893)	Topic 0: Communication	Voice, writing, social, text, media, email, read, sound, talk, image	1110 (22.7%)
	Topic 1: Learning and Education (Translation)	Language, search, learning, video, words, article, translate, work, study, app	1018 (20.8%)
	Topic 2: Software Development	Coding, text, software, programming, python, excel, documentation, web, api, script	1015 (20.7%)
	Topic 3: Learning and Education (Personalized learning)	Research, read, chat, idea, learning, knowledge, system, personal, assistant, conversation	968 (19.8%)
	Topic 4: Marketing and Business (Human resource)	Job, company, writing, idea, business, email, team, support, intelligence, training	782 (16.0%)

The topic distribution varied significantly across platforms. Zhihu users in China focused on content creation and ethical concerns, debating AI's societal impact and its role in replacing jobs. Quora users, primarily from North America, Europe, and Asia, discussed business and marketing applications, particularly AI-driven customer support and communication strategies. Reddit users engaged in a broader range of topics, from AI-enhanced communication tools and software development to its role in personalized learning and education.

3.2 Topic Similarity Analysis

We performed a topic similarity analysis to identify common themes and unique focal points across platforms. Using word embedding models and similarity scoring methods, we quantified and visualized topic similarities. Specifically, we utilized Google's pre-trained Word2Vec model to generate word vectors for topic terms. To ensure cross-platform comparability, Zhihu data was translated into English. Weighted average vectors were calculated based on the word order from the LDA results. Cosine similarity between these vectors produced a cross-platform similarity matrix, visualized as a heatmap (see Fig. 1).

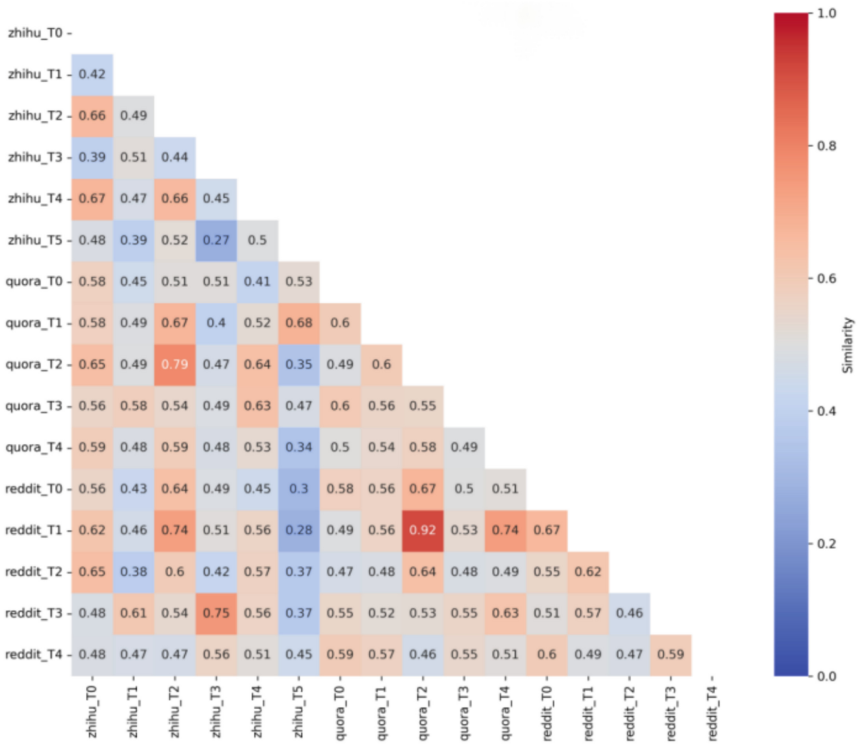


Fig. 1. Cross-Platform topic similarity matrix.

The heatmap reveals high topic similarity across platforms. Zhihu's communication topic (Zhihu_T2) and Quora's communication topic (Quora_T2) show a similarity score of 0.79, highlighting shared discourse on GenAI communication tools for personal and professional use, emphasizing practical applications for efficiency.

Zhihu's learning and education topic (Zhihu_T3) and Reddit's personalized learning topic (Reddit_T3) score 0.75, while Zhihu's communication topic (Zhihu_T2) and Reddit's translation topic (Reddit_T1) score 0.74. These scores indicate mutual interest in GenAI for education, particularly in personalized learning and translation for academic and language purposes.

The highest similarity is between Quora's communication topic (Quora_T2) and Reddit's translation topic (Reddit_T1) at 0.92, reflecting strong alignment in using GenAI for communication and translation. Additionally, Quora's technical learning topic (Quora_T4) and Reddit's translation topic (Reddit_T1) share a similarity score of 0.74, emphasizing the platforms' focus on GenAI's role in learning and communication through educational and translation tools.

3.3 Sentiment Analysis

We used kernel density plots to show the distribution of sentiment values toward GenAI across the three platforms (Fig. 2). A kernel density plot is a smooth curve used to estimate the probability density function of the data, providing an intuitive visualization of the data distribution and its central tendency. The blue line represents the sentiment distribution on Zhihu, the green line represents Quora, and the red line represents Reddit.

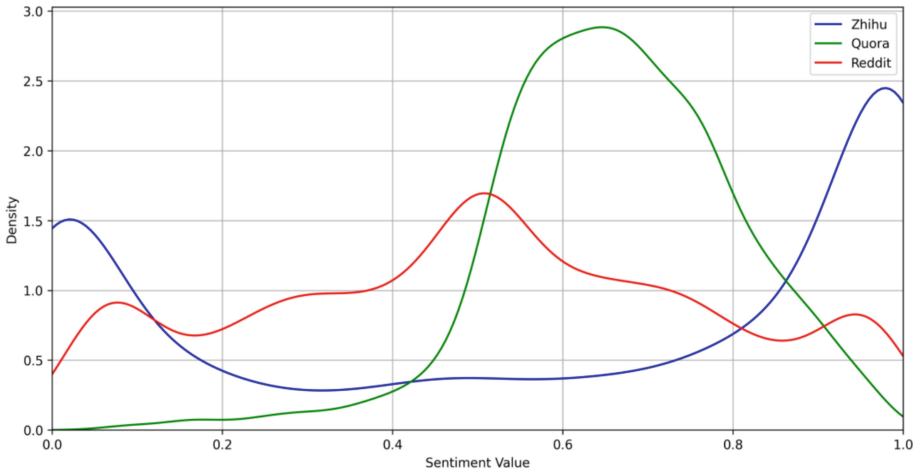


Fig. 2. Comparison of sentiment value distributions across platforms.

The kernel density plot shows that Zhihu's sentiment values exhibit two peaks near 0.1 and 0.9, indicating a strong polarization between very negative and very positive sentiments. Quora's distribution peaks around 0.65, suggesting that most comments are moderately to highly positive. Reddit's sentiment distribution is more balanced, with multiple peaks, including one near 0.5 (neutral sentiment) and some extreme clusters around 0.1 and 0.9.

Overall, Zhihu's discussions are more polarized, Quora's are predominantly positive, and Reddit's sentiments are evenly distributed without a strong tendency toward extremes. Box plots (Figs. 3–5) further illustrate sentiment variations across topics on each platform.

Zhihu's sentiment distribution varies significantly across topics. "Content Creation," "Learning and Education," and "Marketing and Business" topics have similar distributions, with medians around 0.7–0.8. The "Research and Software Development" topic

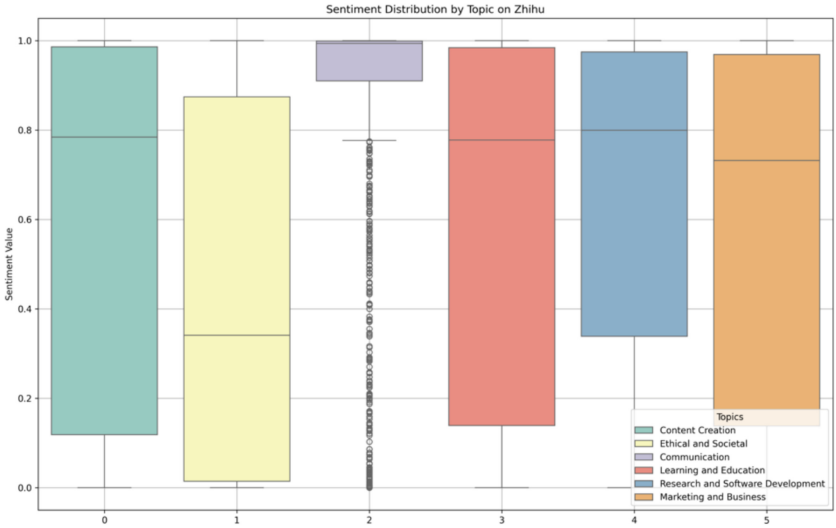


Fig. 3. Sentiment distribution by topic on Zhihu.

ranges from 0.4 to 1.0, with a median of 0.8. "Communication" has the highest median (close to 1.0), reflecting highly positive evaluations, despite some outliers. In contrast, "Ethical and Societal" has a wider range but a median below 0.4, indicating more neutral or negative sentiments.

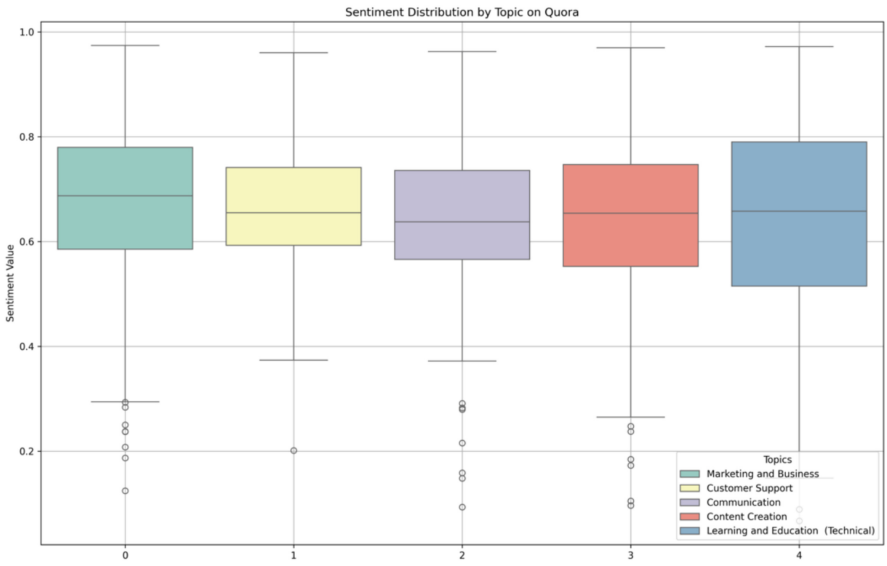


Fig. 4. Sentiment distribution by topic on Quora.

From the box plot of Quora, the median sentiment values for most topics are between 0.6 and 0.7, with the sentiment distribution being relatively concentrated. Most of the boxes are located between 0.5 and 0.7, indicating that users generally hold positive attitudes towards various topics. Each topic has a certain number of outliers, reflecting the extreme views of some users. The sentiment distribution for "Learning and Education (Technical)" is the widest, indicating that Quora users have more emotional fluctuations in this topic.

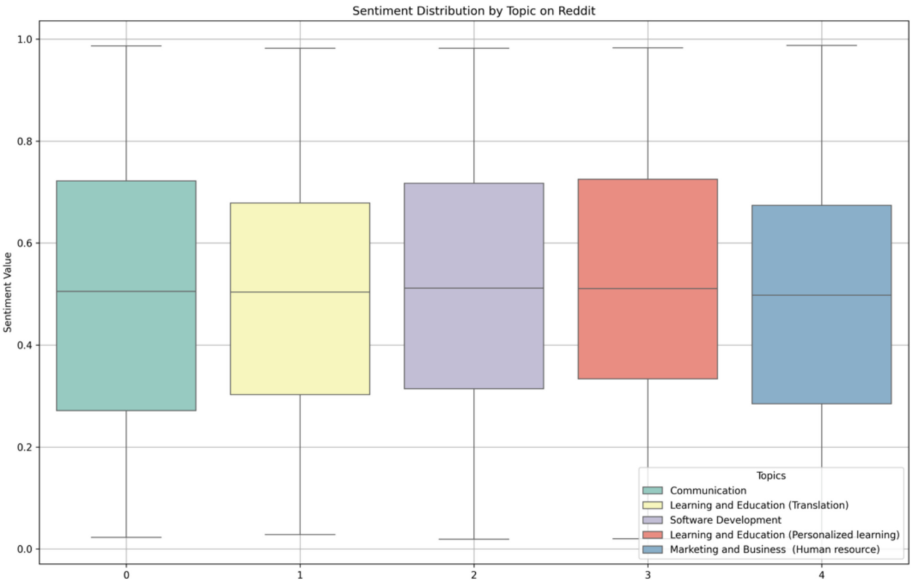


Fig. 5. Sentiment distribution by topic on Reddit.

From the Reddit box plot, the sentiment values across different topics are relatively concentrated. The median sentiment values for most topics are around 0.5, with sentiment values ranging between 0.4 and 0.6. The sentiment distribution is relatively consistent, without any topic showing significant sentiment fluctuations. This indicates that Reddit users' opinions on these topics are relatively balanced.

4 Discussion

The study applied topic modeling, topic similarity analysis, and sentiment analysis to explore GenAI discussions on Zhihu, Quora, and Reddit. LDA identified distinct topics, while similarity analysis highlighted shared themes across platforms. Sentiment analysis further revealed differences in user engagement and perceptions.

Zhihu users focused on content creation and ethical concerns, reflecting its role as a knowledge-sharing platform in China. Quora discussions centered on business and

marketing applications, emphasizing AI-driven customer support and efficiency. Reddit users covered a broader range of topics, from communication tools and software development to learning and education, illustrating GenAI's widespread applicability.

The topic similarity analysis showed cross-platform shared interests in content creation, communication, and education, particularly in personalized learning and translation. Despite this, platform-specific differences emerged, Zhihu users emphasized academic research, Quora users focused on technical learning, and Reddit users explored personalized learning and AI-assisted translation. Similarly, in business discussions, Quora prioritized customer support and marketing strategies, while Reddit discussions revolved around human resources and team collaboration tools.

Sentiment analysis revealed distinct emotional trends. Zhihu discussions were highly polarized, reflecting divided opinions on ethical and practical issues. Quora users exhibited predominantly positive sentiment, indicating optimism about GenAI's business applications. Reddit's sentiment distribution was balanced, reflecting diverse perspectives without extreme views. These findings suggest that platform characteristics and cultural contexts influence how users perceive and discuss GenAI.

5 Conclusion

This study provides insights into GenAI discussions across social Q&A platforms, highlighting both unique and shared themes. The findings suggest that business and marketing tools may be more relevant for Quora users, while learning and research applications better serve Zhihu and Reddit audiences. Addressing Zhihu's polarized sentiments with improved ethical transparency and maintaining Quora's positive sentiment trends could enhance the GenAI user experience.

While this study offers a cross-platform perspective, some limitations exist. Platform selection may affect generalizability, and differences in dataset size across platforms could influence comparisons. Reliance on predefined keywords may exclude relevant discussions. Future research should expand platform coverage, refine sampling strategies, and integrate qualitative insights for a more comprehensive understanding of GenAI engagement.

Acknowledgments. This research is supported by the National Research Foundation, Singapore, under its AI Singapore (AISG) Programme (AISG Award No: AISG3-GV-2023-013). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not reflect the views of AISG.

References

1. Zhao, Y., Wu, L., Zhang, J., Le, T.: How question characteristics impact answer outcomes on social question-and-answer websites. J. Glob. Inf. Manage. (JGIM) **29**(6), 1–21 (2021)
2. Labajová, L.: The state of AI: Exploring the perceptions, credibility, and trustworthiness of the users towards AI-Generated Content (2023)

3. Wu, C., Wang, X., Carroll, J., Rajtmajer, S.: Reacting to generative AI: insights from student and faculty discussions on reddit. In Proceedings of the 16th ACM Web Science Conference (WEBSCI '24), pp. 103–113. Association for Computing Machinery, New York (2024)
4. Wu, Q., Lee, C.S., Goh, D.H.L.: Understanding user-generated questions in social Q&A: a goal-framing approach. J. Am. Soc. Inf. Sci. **74**(8), 990–1009 (2023)
5. Lim, K.K., Lee, C.S.: Review of in-app nudges towards user decision behavior: a topic modeling approach. In: International Conference on Asian Digital Libraries, pp. 263–269 (2023)
6. Wang, H., Wang, J., Zhang, Y., Wang, M., Mao, C.: Optimization of topic recognition model for news texts based on LDA. J. Digital Inf. Manage. **17**(5), 257 (2019)

Generative AI-Driven Personalized Nudges

Kok Khiang Lim(✉) and Chei Sian Lee

Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore
w200004@e.ntu.edu.sg, leecs@ntu.edu.sg

Abstract. Nudging is a concept derived from behavioral economics that subtly influences individuals' behaviors toward beneficial outcomes. In education, nudges have effectively guided students toward positive academic performances. However, nudging faces individual differences and contextual appropriateness constraints, which limit their usefulness in intervening behaviors. To address these limitations, this study leverages insights from behavioral economics and taps into the capabilities of generative artificial intelligence (GenAI) to create personalized nudges delivered via mobile-based web applications. These nudges are tailored to each student's needs and encourage informal learning engagement. A 14-day diary study with 32 university students examined how GenAI analyzed their routine activities and delivered nudges at optimal times. The effectiveness of these nudges was evaluated using latent profile analysis, which identified two student behavior types, with most students responding positively to the learning schedule. Ethical considerations were applied in the nudge design and implementation to minimize disruption to students' routine academic activities, mitigate GenAI biases, and ensure privacy protection. Overall, the results suggested that GenAI-driven nudges are crucial in providing personalized and timely interventions that positively influence students and support their learning. This demonstrates the innovative use of GenAI for educational purposes and its potential to promote personalized learning.

Keywords: Nudge · Generative AI · Digital Technology

1 Introduction

Education has emerged as one of the most promising areas for generative artificial intelligence (GenAI) integration [1, 2], where it offers many innovative ways to support and enhance student learning experiences, from personalized learning platforms to intelligent learning feedback assistants [3, 4]. One of the most intriguing applications of GenAI in education is the integration with nudging - a concept rooted in behavioral science that leverages subtle change in the environment to influence individuals' decisions and behaviors without coercion [5]. The combination of GenAI with behavioral nudging has the potential to profoundly reshape how students engage with their learning environments, presenting new opportunities for educational success.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 359–367, 2025.
https://doi.org/10.1007/978-3-031-94171-9_34

Recent nudge research has shown that personalized nudges offer enhanced outcomes from tailored interventions (e.g., [6, 7]). This is built upon traditional nudges, where issues related to diverse human characteristics and heterogeneous factors are considered to deliver relevant and optimized nudges tailored to individual needs. However, implementing personalized nudges is traditionally manual and resource-intensive, as it involves plowing through vast amounts of data and conducting tedious analysis. For example, in O'Connell and Lang's [7] study, they manually collated, aggregated, and analyzed students' web usage activities and sent emails to nudge students at specific schedules.

Hence, to overcome the limitation and better create personalized nudges, GenAI offers an innovative solution to automatically analyze students' behaviors, progress, and preferences to generate personalized and adaptive nudges that are relevant and context-specific [8]. This study explores using GenAI to generate personalized nudges. The goal is to evaluate the feasibility and the effectiveness of adopting GenAI to create personalized nudges that support students in their informal learning tasks through a 14-day diary study. This required the use of students' non-private and situational data, and such collection is crucial for GenAI to understand students' unique requirements and compute possible nudges. Therefore, the research question is: To what extent does GenAI-driven personalized nudging influence students to engage in informal learning tasks? To address this question, a mobile-based web application is designed and developed to deliver GenAI-driven personalized nudges that facilitate the study.

2 Related Work

2.1 Personalized Nudging

Personalized nudging refers to using behavioral interventions tailored to an individual's characteristics, preferences, and the context of the nudging circumstances to guide people's behavior without coercion or restriction of freedom of choice [5, 9]. This concept is built upon the foundation of traditional nudging, which suggests a subtle change or modification to the choice environment to influence people's behavior toward a desired outcome [5]. It also assumed that all individuals would respond similarly to the nudges presented in the environment, so a one-size-fits-all nudging could resolve most issues. Although positive nudging results have been shown in various domains, such as encouraging good public healthcare behavior [10] and guiding financial decision-making [11], their impact on overall behavioral changes was often mixed due to using standardized nudges to address the diverse circumstances in a similar situation [12, 13]. For instance, the literature review on nudging research in the educational domain has highlighted that students' learning behaviors were heterogeneous and complex (e.g., differences in procrastination tendencies, motivational levels, and learning abilities) even though they were situated in a similar learning environment and that the nudging outcomes were not as desired and expected [12]. This was further evident from the survey study that students' differences also influenced the acceptance of the nudges [14].

Personalized nudging acknowledges this diversity and seeks to address the issue of one-size-fits-all approaches in nudging design [13]. Several studies have explored the potential of personalized nudging in education and have achieved desirable results in

enhancing nudge effectiveness. Dart and Spratt [6] nudged a diverse group of students to improve their learning experiences and outcomes by using personalized emails that included students' names and with an engaging conversational tone. Similarly, Smith et al. [15] sent tailored emails featuring students' names and grades to encourage them to keep up with their homework performance. Additionally, O'Connell and Lang [7] tracked students' interactions with the web-based learning platform. They crafted personalized emails to motivate students to improve their study consistency and learning outcomes based on their content usage and interactions during the course [7]. These findings suggest that personalized nudges effectively enhance behavioral outcomes by addressing the diverse yet unique factors influencing individuals' behaviors, leading to more effective and contextually appropriate interventions. Hence, drawing from the findings of past studies, this study seeks to introduce technology to further optimize personalized nudges by utilizing students' preferences and supporting their learning.

2.2 AI-Nudging

The advancement in AI technology presents a timely opportunity to revolutionize the design and incorporation of nudge interventions. AI can enhance the personalization of nudges by leveraging vast amounts of data and recognizing patterns to tailor interventions based on individual behaviors, preferences, and environmental contexts [16, 17]. These nudges could also be dynamically adjusted based on individual responses and changing circumstances to optimize their impact [16, 17]. Another key benefit of incorporating AI into personalized nudging is its ability to scale by eliminating the manual processes involved in traditional personalized interventions, which are often time-consuming and resource-intensive. Several studies have explored the potential of AI-nudging across different fields, such as marketing [18], healthcare [19], and education [20]. Their findings demonstrated the feasibility of integrating AI and nudging to deliver the desired interventions. However, as AI technology advances, notably GenAI, there is a gap in understanding and evaluating the effect of GenAI and its use in generating and applying personalized nudges on shaping behavior in the educational domain. Hence, this study seeks to evaluate the feasibility and efficacy of GenAI-driven nudges in guiding students' learning.

However, integrating GenAI and nudging is not without challenges. It introduces significant ethical and societal issues that must be approached with caution. Concerns about GenAI model biases, personal privacy, and the potential for manipulation emerge when GenAI systems decide what and when nudges are delivered to people [21, 22]. These concerns become more significant, especially in the educational context, where there is an added responsibility to ensure that any nudge interventions used in education do not conflict with core educational principles or adversely affect other academic areas [23]. Therefore, when designing GenAI-driven nudges, this study will be guided by the nudging principles [5], guidelines in educational nudges [23], and best practices of crafting input prompts for use in GenAI to generate output to develop personalized nudges [24].

3 Method

3.1 Data Collection

This study is part of the larger study which received approval from the Institutional Review Board (IRB-2024-561) to conduct a 14-day diary study at a local university to observe and evaluate students' responses to GenAI-driven personalized nudges delivered through a mobile-based web application (app) during informal learning. A sample of 32 students (five males and 27 females; 26 full-time and six part-time) aged between 19 and 47 years ($M = 27.6$, $SD = 5.48$) responded and completed the 14-day study, and were compensated for their time and effort upon completion.

3.2 Nudge Implementation

A web app was designed and developed to deliver GenAI-driven personalized nudges, subtly delivered through reminder notifications on upcoming learning tasks, default task schedules, and learning advice, that were optimized according to the student's daily routine activities. Besides providing the default GenAI-generated schedule, the app allows the students to interact and adjust the schedule further to give students the necessary autonomy and control over the nudges.

The app's conceptualization and prototyping were made in Figma[1] to design the app's graphical user interface and test the logical flow. The finalized app was developed and hosted on Bubble.io[2], a visual programming platform that provides the development of web applications without writing programming codes [25]. The app consisted of an interactive graphical user interface (GUI) for students to interact with the nudges supported by backend modules that included interfacing with the GenAI server through OpenAI's API[3], databases, and a nudging scheduler.

The students installed the web app on their mobile devices during the study and logged in using the assigned and anonymized credentials. The student first entered three daily routine activities - school, dinner, and sleep time, and the app used this information to recommend a learning schedule and offer helpful learning advice. Nudges were triggered at scheduled times to notify students about their learning. Students' interaction and usage activities within the app, such as modifying learning schedules and responding to nudge notifications, were recorded and used in the post-study analysis to evaluate nudge effectiveness.

3.3 Measurement and Analysis of Nudge Effectiveness

Nudge effectiveness refers to how well a nudge has successfully influenced or steered individuals to make decisional changes aligned with the desired outcome without using forbidding options or coercive means [5]. In this study, nudge effectiveness is measured by the time the student takes to act upon the nudge intervention's trigger, derived from

[1] https://www.figma.com/

[2] https://bubble.io/

[3] https://platform-openai.com/docs/guides/gpt/chat-completions-api

when the scheduled nudges were triggered to remind students to commence the tasks until the students responded to the nudging and attempted learning.

The measurements were analyzed using latent profile analysis (LPA) to determine the degree to which students had responded to the nudges [26]. LPA is a model-based probabilistic clustering analysis that classifies sub-groups of people who share similar responses within a group with different response patterns [27]. The number of sub-groups is determined through modelling to obtain the best-fit results by assessing the Bayesian Information Criterion (BIC) [28], Akaike Information Criterion (AIC) [29], and entropy as a measure to evaluate the separation between classes [30]. A lower BIC and AIC values and an entropy closer to 1 represent a better fit and representation of classes [30, 31]. The LPA is performed using the R statistical tool with tidyLPA package [32]. Finally, the computed sub-groups were given descriptive labels.

3.4 Ethical Considerations

The implementation of GenAI-driven personalized nudges was carefully planned and designed, guided by the principles of nudges [5] and suggested guidelines for applying nudges in education [23]. This ensured that the overall nudging intervention would not disrupt students' routine academic activities or induce the risks of causing undesirable or unwanted learning behavior [23]. In addition, throughout the nudge intervention, students retained control over their decisions in taking up the learning, thus upholding the principles of nudging that are non-forceful and non-coercive [5].

To mitigate GenAI biases, the input prompts to the GenAI servers were meticulously crafted and tested extensively to ensure that the GenAI-generated output was coherent and reliable within the learning context [24]. The output was also checked for consistency in generating credible task schedules before being deployed in this study. Further, to ensure privacy protection, no personal identifier or sensitive information was crafted or used to develop the input prompts.

4 Results

Table 1 shows the LPA modeling results. LPA models with 1 to 5 classes were executed to identify the best model fit. The findings indicated that model 6, which revealed a 2-class solution, provided the best fit for the observed data. That is, the 2-class solution adequately differentiated the different types of students who respond to nudges.

A density plot (see Fig. 1), generated using the selected LPA model 6 and 2-class parameters, clearly illustrates a separation between the classes. Class 1 occurred around $t = 0$, indicating that students responded shortly after the nudge intervention was triggered. Meanwhile, class 2 shows a deviation from $t = 0$, suggesting some students delayed their follow-up action on informal learning. As such, classes 1 and 2 were labeled as 'normal response' and 'delay response,' respectively.

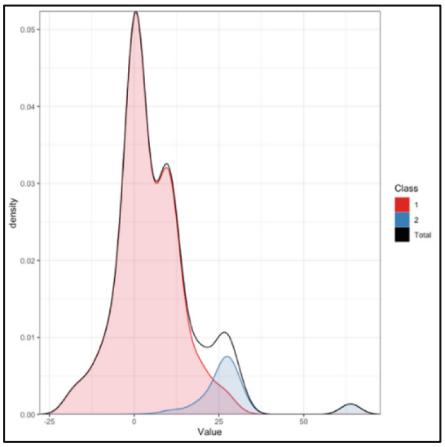
5 Discussion

This study evaluated the effectiveness of using GenAI-driven personalized nudges to guide students' learning behavior over a 14-day diary study. The results showed that most students responded positively to the GenAI-driven personalized nudges during informal

Table 1. LPA model fit statistics for the nudging response behavior.

Model	Classes	AIC	BIC	Entropy
1	1	1613.542	1623.963	1
1	2	1549.262	1567.498	0.904
1	3	1508.967	1525.018	0.907
1	4	1442.051	1475.918	0.944
1	5	1417.425	1459.107	0.947
6	1	1358.098	1371.124	1
6	**2**	**1200.478**	**1229.135**	**0.980**
6	3	1201.132	1245.420	0.642
6	4	1195.543	1255.462	0.773
6	5	1181.909	1257.459	0.963

Note. Model 1 = Equal variances, covariances; Model 6 = Unequal variances, unequal covariances; Bold font shows the selected model and class



Note. Class 1 = Normal response, Class 2 = Delay response.

Fig. 1. LPA density plot for students' nudging response behavior.

learning. This was evaluated using the LPA, which had identified two distinct students' responses to the nudges: one that had a significantly higher proportion of students' responses happened within the immediate nudge influences (termed 'normal response') and the other with a smaller number that showed delays in their responses (termed 'delayed response). This finding further supports the notion that to meet the diverse population, nudges should be personalized, taking account of individual characteristics, preferences, and the context of the nudging environment to optimize nudge techniques in enhancing their efficacy [12, 33, 34]. In addition, it demonstrated GenAI's ability to

comprehend students' needs and provide appropriate and relevant responses to support students' learning.

Although personalized nudges tend to be more effective, in contrast, this is not always the case. The findings showed that delayed responses were also present. According to Ingendahl et al. [35], personality traits, such as NFU ('need for uniqueness' where individuals differentiate themselves from others) and NFC ('need for cognition' where people tend to take part in high cognitive load tasks), may not be influenced by specific nudging techniques, such as defaults and social influence. The study by Lim and Lee [14] further revealed that differences in student characteristics would also affect how they perceived and accepted nudges to support their learning.

6 Conclusion

Overall, the findings provided a fresh perspective on the innovative use of GenAI for educational purposes, demonstrating GenAI's potential to create effective personalized nudges tailored to individual students' needs in supporting their learning. This is particularly critical as past studies have highlighted that no one-size-fits-all nudges can fulfill all scenarios or demographics [12, 33, 34]. The findings also revealed students' nudge-respond behavior, which may help improve the nudging technique design. Further, this study is guided by nudge principles and guidelines and adopted the best practices in crafting the GenAI's input prompt to ensure that ethical concerns are mitigated or reduced to a minimum. As GenAI is still evolving, future studies implementing GenAI-driven interventions to influence students' behavior should keep current on the ethical guidelines or adhere to policies governed by oversight bodies [22].

Two limitations were observed in this study. Firstly, the sample size leaned towards full-time female students within the students' population, and this may skew the results because of the broadly similar students' demographic, which moderated the way they react to nudges. Past studies have shown that demographic differences, such as academic discipline and student cohorts [14, 36], tend to exhibit differing results. As such, the findings from this study could not be generalized across other population segments. Nonetheless, the findings from this study could extend to future studies and provide a preliminary understanding of the effects and success of using GenAI-driven nudges to support learning. In addition, future studies could expand to a diverse student population in understanding the effects of GenAI-driven personalized nudging.

Secondly, this study used students' nudge responses to objectively evaluate nudge effectiveness using LPA. Although the results provided a deterministic outcome, it could not capture the full context of student behavior and learning to elucidate the outcomes. Therefore, future studies may consider incorporating other measurement instruments to assess students' underlying factors influencing their behaviors. Despite these limitations, the present study yields critical insights into Gen-AI-driven nudges on students.

Acknowledgments. This research is supported by the Ministry of Education, Singapore under its Tertiary Education Research Fund (Award MOE2023-TRF-028). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Ministry of Education, Singapore. The authors also thank How Jun Yao Jonathan, Moo Kee Khong and Sun Shaolei, for assisting in the data collection.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. World Economic Forum. https://www.weforum.org/stories/2024/04/future-learning-ai-revolutionizing-education-4-0/
2. GovTech. https://www.tech.gov.sg/media/technews/ai-in-education-transforming-singapore-education-system-with-student-learning-space/
3. Chan, G.: askST, What are MOE's newest artificial intelligence tools, and how are schools using them? The Straits Times. Singapore Press Holdings, Singapore, 29 July 2025
4. https://onlinedegrees.sandiego.edu/artificial-intelligence-education/
5. Thaler, R.H., Sunstein, C.R.: Nudge: Improving decisions about health, wealth, and happiness. Yale University Press (2008)
6. Dart, S., Spratt, B.: Personalised emails in first-year mathematics: exploring a scalable strategy for improving student experiences and outcomes. Student Success **11**, 1–12 (2020). https://doi.org/10.5204/ssj.1543
7. O'Connell, S.D., Lang, G.: Can personalized nudges improve learning in hybrid classes? experimental evidence from an introductory undergraduate course. J. Res. Technol. Educ. **50**, 105–119 (2018). https://doi.org/10.1080/15391523.2017.1408438
8. Harrison, R.M., Lapteva, E., Bibin, A.: Behavioral nudging with generative AI for content development in SMS health care interventions: case study. JMIR AI **3**, e52974 (2024). https://doi.org/10.2196/52974
9. Karlsen, R., Andersen, A.: Recommendations with a Nudge. Technologies 7, (2019). https://doi.org/10.3390/technologies7020045
10. Lee, C.S., Lim, K.K., Kim, H.K.: Nudging public health behaviors to prevent COVID-19: A systematic review. Health Communication 1–12 (2024). https://doi.org/10.1080/10410236.2024.2317567
11. Thaler, R.H., Benartzi, S.: Save more tomorrow: Using behavioral economics to increase employee saving. J. Political Econ. **112**, 164–187 (2004). https://doi.org/10.1086/380085
12. Lim, K.K., Lee, C.S.: Nudging learning behaviour: a systematic review. Proc. Assoc. Inf. Sci. Technol. **59**, 744–746 (2022). https://doi.org/10.1002/pra2.712
13. Mills, S.: Personalized nudging. Behavioural. Public Policy **6**, 150–159 (2022). https://doi.org/10.1017/bpp.2020.7
14. Lim, K.K., Lee, C.S.: Acceptability of nudge in digital learning environment. Proc. Assoc. Inf. Sci. Technol. **61**, 564–569 (2024). https://doi.org/10.1002/pra2.1059
15. Smith, B.O., White, D.R., Kuzyk, P.C., Tierney, J.E.: Improved grade outcomes with an e-mailed "grade nudge." J. Econ. Educ. **49**, 1–7 (2018). https://doi.org/10.1080/00220485.2017.1397570
16. Nyman, S.: The birth of AI-driven nudges. In: 56th Hawaii International Conference on System Sciences, pp. 5252–5261. (Year). 10125/103276
17. Sadeghian, A.H., Otarkhani, A.: Data-driven digital nudging: a systematic literature review and future agenda. Behav. Inf. Technol. **43**, 3834–3862 (2024). https://doi.org/10.1080/0144929X.2023.2286535
18. Gao, Y., Liu, H.: Artificial intelligence-enabled personalization in interactive marketing: a customer journey perspective. J. Res. Interactive Mark., 1–18 (2022). https://doi.org/10.1108/JRIM-01-2022-0023

19. Joachim, S., Forkan, A.R.M., Jayaraman, P.P., Morshed, A., Wickramasinghe, N.: A nudge-inspired AI-driven health platform for self-management of diabetes. Sensors **22**, 4620 (2022). https://doi.org/10.3390/s22124620
20. Rodriguez, M.E., Guerrero-Roldán, A.E., Baneres, D., Karadeniz, A.: An intelligent nudging system to guide online learners. Int. Rev. Res. Open Distributed Learn. **23**, 41–62 (2022). https://doi.org/10.19173/irrodl.v22i4.5407
21. Dalecke, S., Karlsen, R.: Designing dynamic and personalized nudges. In: Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics, pp. 139–148. Association for Computing Machinery, Biarritz, France (2020)
22. Schmauder, C., Karpus, J., Moll, M., Bahrami, B., Deroy, O.: Algorithmic nudging: the need for an interdisciplinary oversight. Topoi (2023). https://doi.org/10.1007/s11245-023-09907-4
23. Weijers, R.J., de Koning, B.B., Paas, F.: Nudging in education: from theory towards guidelines for successful implementation. Eur. J. Psychol. Educ. **36**, 883–902 (2021). https://doi.org/10.1007/s10212-020-00495-0
24. Marvin, G., Hellen, N., Jjingo, D., Nakatumba-Nabende, J.: Prompt engineering in large language models. Data Intelligence and Cognitive Informatics 387–402 (2024). https://doi.org/10.1007/978-981-99-7962-2_30
25. https://bubble.io/
26. Vangsness, L., Young, M.E.: Turtle, task ninja, or time waster? Who cares? Traditional task-completion strategies are overrated. Psychol. Sci. **31**, 306–315 (2020). https://doi.org/10.1177/0956797619901267
27. Mathew, A., Doorenbos, A.Z.: Latent profile analysis – an emerging advanced statistical approach to subgroup identification. Indian J. Continuing Nursing Educ. **23** (2022)
28. Schwarz, G.: Estimating the dimension of a model. Ann. Stat. **6**, 461–464 (1978)
29. Akaike, H.: Factor analysis and AIC. Psychometrika **52**, 317–332 (1987). https://doi.org/10.1007/BF02294359
30. Celeux, G., Soromenho, G.: An entropy criterion for assessing the number of clusters in a mixture model. J. Classif. **13**, 195–212 (1996). https://doi.org/10.1007/BF01246098
31. Tein, J.Y., Coxe, S., Cham, H.: Statistical power to detect the correct number of classes in latent profile analysis. Struct. Equ. Model. **20**, 640–657 (2013). https://doi.org/10.1080/10705511.2013.824781
32. Rosenberg, J., Beymer, P., Anderson, D., Schmidt, J.: TidyLPA: an R package to easily carry out latent profile analysis (LPA) using open-source or commercial software. J. Open Source Softw. **3**, 978 (2018). https://doi.org/10.21105/joss.00978
33. Kawa, C., Gijselaers, W.H., Nijhuis, J.F.H., Ianiro-Dahm, P.M.: Are you "nudgeable"? factors affecting the acceptance of healthy eating nudges in a cafeteria setting. Int. J. Environ. Res. Public Health **19**, 4107 (2022). https://doi.org/10.3390/ijerph19074107
34. Schneider, C., Weinmann, M., vom Brocke, J.: Digital nudging: guiding online user choices through interface design. Commun. ACM **61**, 67–73 (2018). https://doi.org/10.1145/3213765
35. Ingendahl, M., Hummel, D., Maedche, A., Vogel, T.: Who can be nudged? examining nudging effectiveness in the context of need for cognition and need for uniqueness. J. Consum. Behav. **20**, 324–336 (2021). https://doi.org/10.1002/cb.1861
36. Chen, Y.H., Lin, Y.J.: Validation of the short self-regulation questionnaire for Taiwanese college students (TSSRQ). Front. Psychol. **9**, 259 (2018). https://doi.org/10.3389/fpsyg.2018.00259

User-Perceived Quality of GenAI Chatbots in a Business Context – A Construct Definition

Milad Morad(✉) and Britta Essing

Schloss Birlinghoven, Fraunhofer FIT, Sankt Augustin, Germany
milad.morad@fit.fraunhofer.de

Abstract. Generative Artificial Intelligence (GenAI) is increasingly integrated into business environments with the expectation of enhancing productivity across various organizational levels. However, while technical implementation is prioritized, critical aspects such as employee-perceived quality remain underexplored. Existing usability specifications, primarily developed for deterministic systems, fail to capture GenAI's adaptive and probabilistic nature, necessitating a redefinition of user-perceived quality to establish valid design requirements and measurement instruments.

This study aims to conceptualize a user-centered quality framework for GenAI chatbots in a business context. Employing a mixed-method approach, the research includes 12 semi-structured interviews and a survey of 512 respondents, followed by an exploratory factor analysis to identify key quality dimensions. The findings reveal five principal dimensions of user-perceived quality: (1) Productivity of Responses, (2) Trustworthiness, (3) Contextual Fit, (4) Domain Suitability, and (5) Operational Usability. These dimensions provide a structured perspective on how employees evaluate GenAI chatbots, addressing critical gaps in existing usability models.

The study contributes to the Human-Computer Interaction (HCI) field by extending traditional usability and user experience (UX) frameworks to accommodate the dynamic nature of GenAI. The identified quality dimensions offer organizations a structured approach to assessing and refining GenAI solutions based on user feedback. Future research will focus on further empirical validation through a large-scale study, investigating the correlation between quality dimensions, user satisfaction, and long-term GenAI adoption in business settings. By establishing a robust evaluation framework, this research supports the development of GenAI chatbots that enhance both employee satisfaction and organizational performance.

Keywords: Generative AI · User-perceived Quality · Human-Centered AI

1 Introduction

Generative Artificial Intelligence (GenAI) has gained attention across industries due to its ability to automate knowledge work [1]. Its integration into business environments as enterprise chatbots is driven by the expectation that it can improve productivity at multiple levels, from individual employees to entire organizations [2].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 368–376, 2025.
https://doi.org/10.1007/978-3-031-94171-9_35

However, companies predominantly focus on the technical realization and implementation of GenAI chatbots, while critical aspects such as employee-centricity remain underexplored [3]. Integrating technological excellence and human-centered design is imperative to enable valuable human-machine collaboration [4]. Ultimately, the employees or customers must accept and integrate GenAI solutions into their daily work.

While user-centered design has a long tradition in HCI, and usability – such as defined in ISO 9241-11 – has evolved into a well-developed and validated quality framework, it remains inadequate for designing and evaluating interactive GenAI systems. GenAI is an adaptive system unlike traditional information systems that operate on rule-based functions [5]. It learns from user interactions, adapts to new requirements, and generates content dynamically. While this versatility enhances its applicability, it also introduces new challenges in evaluating its impact [6].

The first challenge that needs to be overcome is to redefine the construct of the user-centered quality view beyond existing usability concepts to have suitable and valid design requirements for GenAI systems and thus provide the basis for future measurement instruments.

A key challenge lies in methodologically assessing how GenAI is utilized within corporate environments, spanning diverse roles, competencies, and responsibilities. Understanding in which task contexts GenAI is most effectively used enables organizations to maximize its potential. A structured evaluation of user-perceived quality will help companies identify the most valuable areas where GenAI supports employees and its potential remains underutilized [2].

User satisfaction can be a key metric in evaluating the success of GenAI adoption [7]. Existing measurement instruments from human-computer interaction (HCI) methodologies, such as the System Usability Scale [8] and the User Experience Questionnaire [9], were primarily developed for rule-based software systems [6]. Given GenAI's different nature, these traditional metrics are insufficient to assess its user-perceived quality and interaction adaptability [10].

This study aims to provide a construct definition of a user-centered quality view of GenAI chatbots, to lay the foundation for companies to assess user-perceived quality and refine GenAI design and functionality for business applications. To achieve this, the author team addresses the following research question (RQ):

- What is the user-perceived quality of GenAI chatbots in a business context?

Methodologically, this study adopts a mixed-method approach that combines qualitative and quantitative research methods. First, the author's team conducted 12 exploratory semi-structured in-depth interviews to uncover initial dimensions of user-perceived quality. A survey involving over 500 participants was administered, followed by an exploratory factor analysis (EFA) to validate and refine the measurement model. The findings provide insights into the key dimensions of user-perceived quality of GenAI chatbots, offering organizations a scientifically based perspective on the quality of their GenAI chatbots from an employee's viewpoint.

2 Related Work

Evaluating software quality, usability, and user experience has been a core focus in human-computer interaction (HCI). Traditional frameworks such as ISO 25010 define software quality through attributes like functional suitability, usability, reliability, and maintainability, providing structured guidelines for assessing system performance in deterministic environments [11]. Similarly, ISO 9241-110 establishes usability principles centered around efficiency, effectiveness, and user satisfaction [12]. These models have been widely applied to rule-based software, where system outputs remain predictable, and interactions follow pre-defined steps.

Alongside these standards, usability heuristics such as Jakob Nielsen's principles focus on factors like consistency, error prevention, and user control, which are critical for designing static user interfaces [13]. Further perspectives, such as Marc Hassenzahl's UX framework, differentiate between pragmatic UX, goal-oriented system efficiency, and hedonic UX, which encompasses emotional engagement and user satisfaction [14]. While these models have successfully guided software design and evaluation, they assume deterministic interactions and structured information retrieval, making their application to GenAI systems challenging [15].

Unlike conventional software, GenAI operates probabilistically, generating context-dependent outputs and adapting dynamically to user input. This introduces several challenges that existing software quality frameworks miss to address. First, traditional models emphasize predictability and expectation conformity, whereas GenAI systems often provide varying responses to the same query, affecting user trust [16]. Second, explainability and transparency are not explicitly integrated into usability standards, yet they are crucial for AI adoption, as users need to understand how responses are generated [17]. Furthermore, error handling in classical usability frameworks assumes systems can recognize and prevent incorrect outputs, whereas GenAI models may generate hallucinations, false or misleading information without an inherent validation mechanism [18].

Given these challenges, existing HCI measurement instruments also struggle to assess user-perceived quality of GenAI chatbots. The System Usability Scale (SUS) remains widely used for evaluating usability in structured interfaces, yet it does not account for GenAI characteristics like adaptability or transparency [8]. Similarly, the User Experience Questionnaire (UEQ) focuses on pragmatic and hedonic UX but assumes predictable system behavior, making it insufficient for assessing GenAI interactions [9].

The limitations of existing models identified underscore the necessity of conceptualizing user-perceived quality of GenAI chatbots. This study seeks to develop a structured understanding of how users assess the quality of GenAI by systematically analyzing the key dimensions that shape their evaluation, particularly in a business context. By identifying and defining these dimensions, this research provides a theoretical foundation for future studies to develop measurement approaches specifically tailored to the unique challenges posed by the adaptive, probabilistic, and generative nature of GenAI interactions.

3 Methodology

To address the RQ, the author team conducted a mixed-method approach combining qualitative and quantitative methods to identify and validate key quality dimensions of GenAI chatbots. To derive initial insights, 12 in-depth interviews with expert and non-expert users of GenAI systems were conducted using a semi-structured guideline. ChatGPT is the most frequently used GenAI chatbot among the interviewees, followed by in-house chatbots, Microsoft Copilot, and GitHub Copilot. The primary user intent categories reported by the interviewees are content creation and Q&A applications, followed by knowledge searching, learning, and leisure activities. The interview explored the interviewees' general understanding of quality, as well as the characteristics and examples of both essential and less critical traits of GenAI chatbots. Following an inductive approach, the author team conducted a qualitative content analysis based on Mayring's framework [19] to categorize and structure the statements. This process resulted in identifying 27 distinct quality dimensions, ranked by their frequency of occurrence.

Subsequently, the team of authors pursued data reduction, for which questionnaire items were first developed that comprehensively covered the identified dimensions, with each dimension represented by at least three items, in addition to comparing the developed items with the established measurement instruments, a content validity test [20] was carried out. The author team conducted a content validity test (CVI) with AI experts to assess the content validity of the developed questionnaire items. Each item was rated on a 4-point scale regarding its relevance and comprehensibility. The CVI score was calculated as the proportion of items rated as relevant (score ≥ 3) by most experts. Based on this analysis, the questionnaire was refined by reducing the number of items.

In the subsequent stage, a pre-study was conducted, and 512 people were interviewed using the developed questionnaire. The pre-study included 512 randomly selected respondents from Europe and North America, all of whom reported using GenAI in a business context at least once per week. Each item was rated on a 7-point Likert scale, where participants indicated the importance of a GenAI chatbot possessing the respective characteristic. The high number of response options per item was chosen to mitigate ceiling effects by allowing for more significant differentiation in participant responses.

After gathering data, a data cleaning and preprocessing procedure was carried out to remove inconsistencies, handle missing values, and ensure the integrity of the dataset. This procedure resulted in 502 responses being deemed valid for the analysis, which was then conducted using SPSS Statistics.

To identify the underlying dimensions of user perception of GenAI, an exploratory factor analysis (EFA) was performed using the maximum likelihood model [21]. The number of factors was determined using various statistical criteria, including the Kaiser criterion [22], the Cattell criterion [23], and a parallel analysis [24]. The factor analysis was conducted iteratively to identify a robust factor structure, minimize cross-loadings, and improve model interpretability.

The quality of the developed model was examined by evaluating several criteria. The Kaiser-Meyer-Olkin (KMO) test yielded a value of 0.92, indicating an excellent level of sampling adequacy for factor analysis [25]. The Bartlett's test for sphericity was

significant with $p < 0.001$, confirming that the correlation matrix was suitable for factor extraction and did not resemble an identity matrix [26].
Iterative optimization was performed to achieve factor loadings above 0.40. The Cronbach's alpha coefficient was calculated to assess internal consistency, yielding values of $\alpha > 0.70$ for all extracted factors, indicating a high reliability of the scales. As part of the internal consistency assessment, items and categories with low or weak internal consistency correlations were identified and systematically removed or merged, resulting in five key factors. This refinement process ensured that each retained item contributed meaningfully to the model's overall construct validity and reliability.
While the pre-study findings offer initial insights and a preliminary factor structure, a subsequent main study is required. The main study will involve a large sample of company users to assess the identified dimensions' stability, reliability, and generalizability [27]. The external validity will also be evaluated by comparing the model with established measurement frameworks. The results of this validation will aid in developing the measurement instrument and ensure its applicability across various business contexts.

4 Results

An exploratory factor analysis (EFA) was conducted to ascertain the central quality dimensions of GenAI chatbots. The study produced a strong factor structure of five main dimensions that can be viewed as key dimensions of user-perceived quality in GenAI chatbots within a business context. These dimensions capture different aspects of the user experience and provide information on how well GenAI systems meet user requirements. The following sections present each of these dimensions in detail.

- **Productivity of responses**. This dimension evaluates the clarity, precision, and reliability of the responses generated by GenAI chatbots. A high level of productivity ensures that information is factually accurate, well-structured, and easily usable for professional tasks. Completeness, precision, and reliability of information are essential, as incomplete or ambiguous outputs hinder user efficiency. A GenAI chatbot must produce specific and contextually appropriate responses that align with user queries to be deemed highly productive. Information should be organized logically and presented clearly, allowing users to process and apply the content in their workflows efficiently. In summary, the productivity of responses reflects a GenAI chatbot's capability to provide coherent, complete, and precise information that meets professional standards.
- **Trustworthiness**. This dimension assesses the transparency, traceability, and verifiability of responses provided by GenAI chatbots. Trust is critical in AI adoption, especially in business environments where decisions must rely on accurate and justifiable information. A core aspect of trustworthiness is the verifiability of responses, allowing users to evaluate the correctness of the provided content independently. GenAI chatbots should supply references, citations, or validation mechanisms to enable users to trace the origins of information and verify outputs, thereby enhancing confidence and reducing uncertainty. Reliable GenAI systems integrate error detection and mitigation, ensuring that responses remain robust while minimizing the risk of misinformation. In summary, trustworthiness is a fundamental dimension of

user-perceived quality, which directly influences user confidence and the system's integration into essential business processes.

- **Contextual Fit**. This dimension assesses the GenAI chatbot's ability to adapt to specific user contexts, ensuring that generated responses remain relevant and applicable across various work environments. A crucial element of contextual fit is flexibility, meaning the system must adjust to different tasks, technical requirements, and evolving user needs. Another essential aspect is personalization, which refers to the system's capacity to integrate user feedback and refine responses as necessary. An effective GenAI chatbot learns from interactions, enhancing contextual relevance over time. Reliability is also vital; users must be able to rely on the system to comprehend their work context and provide meaningful support, even when they lack domain-specific knowledge. In summary, contextual fit reflects a GenAI system's flexibility, personalization, and reliability, ensuring it effectively adapts to different usage scenarios, incorporates user input, and delivers contextually appropriate assistance.
- **Domain Suitability**. This dimension assesses a GenAI's ability to accurately apply domain-specific knowledge, ensuring that generated outputs meet the specialized requirements of professional fields. Domain suitability involves the precise application of industry-specific standards and frameworks, guaranteeing terminological accuracy and contextually appropriate content to prevent misinterpretations. Furthermore, the relevance of responses for domain-specific applications influences the system's practical utility; therefore, GenAI must be capable of adapting its outputs to specialized tasks and ensuring that generated insights, recommendations, or solutions align with the expectations of employees within a given field. In summary, domain suitability reflects a GenAI system's ability to apply industry-specific knowledge, utilize precise terminology, and deliver relevant outputs for domain-specific applications.
- **Operational Usability**. This dimension evaluates the ease of use and adaptability of GenAI chatbots in operational environments. Operational usability focuses on intuitiveness, seamless interaction without extensive training, and effective workflow integration. The system should be a supportive and flexible tool to maximize utility, aligning with established processes instead of necessitating significant operational changes. Flexibility enhances usability by allowing the system to modify its responses based on user interactions and changing task requirements. Operational usability embodies a GenAI system's intuitiveness, workflow integration, adaptability, and interactive features. It is crucial for ensuring user efficiency and overall system satisfaction in business settings.

The identified quality dimensions – Productivity of Responses, Trustworthiness, Contextual Fit, Domain Suitability, and Operational Usability – provide a structured approach for evaluating GenAI chatbots within a business context. Each dimension is operationalized through specific questionnaire items, ensuring a thorough system performance assessment from a user perspective. Based on these dimensions, a validated measurement instrument should be developed in the subsequent steps.

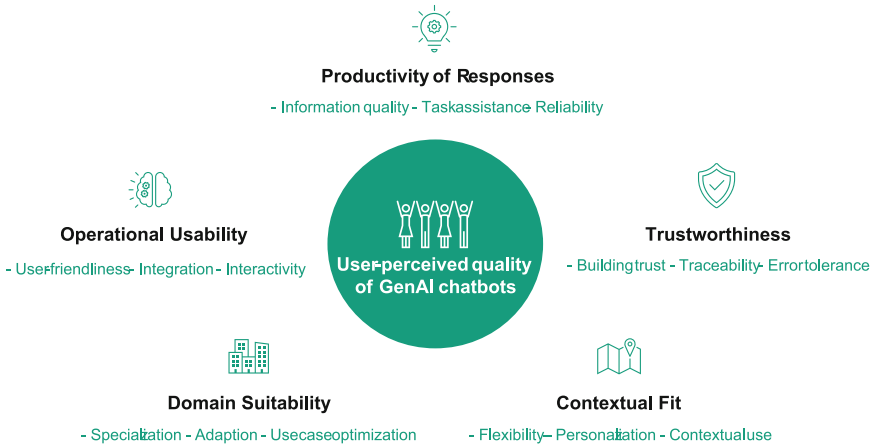


Fig. 1. Dimensions of the user-perceived quality of GenAI chatbots in a business context.

5 Conclusion

This study investigates the research question of what defines the user-perceived quality of GenAI chatbots within a business context. An exploratory factor analysis revealed five key quality dimensions: Productivity of Responses, Trustworthiness, Contextual Fit, Domain Suitability, and Operational Usability. These dimensions effectively capture essential aspects that significantly impact users' evaluations of GenAI.

Unlike traditional usability and UX models, which were primarily designed for deterministic software, the identified quality dimensions address the challenges presented by adaptive, probabilistic, and generative systems. These established models do not sufficiently represent GenAI's key characteristics.

The findings of this study contribute to scientific knowledge by enhancing the existing literature with insights on quality specifically tailored to GenAI. The identified quality dimensions enable companies to systematically evaluate GenAI chatbots based on user-side quality for the first time. This facilitates informed decisions regarding introducing, implementing, and further developing these systems while clarifying the areas where GenAI chatbots in a business context offer the most significant benefits or still need further enhancement.

The study is limited by its exploratory nature and the preliminary phase, which includes a pre-study. While the quality dimensions identified offer valuable insights, they need further empirical validation. The following research stage plans a comprehensive main study with a larger, representative sample of company employees. It will also explore the connection between quality dimensions and user satisfaction, refining and validating the measurement tool. Analyzing how these dimensions impact long-term GenAI adoption, user satisfaction, and task performance across various business contexts is a promising direction for future research. Ultimately, this research aims to support developing and implementing GenAI chatbots that better meet user needs, boosting employee productivity and satisfaction.

Acknowledgments. The icons used in Fig. 1 were obtained from Microsoft PowerPoint (Microsoft 365), and they are used in accordance with the Microsoft Office licensing terms. In writing this paper, the authors utilized Grammarly (Ver. 1.2.142.1621), ChatGPT (Ver. 1.2025.41), DeepL (Ver. 4.4.2.7961), and Microsoft Word to enhance linguistic clarity and presentation. Full responsibility for the contents lies with the authors.

References

1. Schmager, S., Pappas, I.O.: Understanding Human-Centered AI: a review of its defining elements and a research agenda. Behav. Inf. Technol. (2025). https://doi.org/10.1080/0144929X.2024.2448719
2. Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., Kluge, A.: Defining human-AI teaming the human-centered way: a scoping review and network analysis. Front. Artif. Intell. 6, Article 1250725 (2023). https://doi.org/10.3389/frai.2023.1250725
3. Molka-Danielsen, J., Rasool, J., Smith, C.H.: Design and deployment considerations for ethically advanced technologies for human flourishing in the workplace. In: Bhutkar, G., et al. Human Work Interaction Design. Artificial Intelligence and Designing for a Positive Work Experience in a Low Desire Society. HWID 2021. IFIP AICT, vol 609, pp. 101–118. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-02904-2_5
4. Xu, W., Dainoff, M.J., Ge, L., Gao, Z.: From human-computer interaction to human-AI interaction: new challenges and opportunities for enabling human-centered AI. arXiv preprint arXiv:2105.05424 (2021). https://arxiv.org/abs/2105.05424
5. Hendriks, P., Sturm, T., Geis, M., Grimminger, T., Mast, B.: Exploring collaboration in human-artificial intelligence teams: a design science approach to team-AI collaboration systems. In: PACIS 2024 Proceedings, Track 08. AIS, Hong Kong (2024). https://aisel.aisnet.org/pacis2024/track08_digtech_fow/track08_digtech_fow/1
6. Lindgren, H.: Emerging Roles and Relationships Among Humans and Interactive AI Systems. Int. J. Hum. –Comput. Interact., 1–23 (2024). https://doi.org/10.1080/10447318.2024.2435693
7. Schmutz, J.B., Kerstan, S., Georganta, E.: AI-Teaming: Redefining Collaboration in the Digital Era. Curr. Opin. Psychol. **58**, Article 101837 (2024). https://doi.org/10.1016/j.copsyc.2024.101837
8. Brooke, J.: SUS: A 'Quick and Dirty' usability scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B. (eds.) Usability Evaluation in Industry, pp. 189–194. CRC Press, London (1996). https://doi.org/10.1201/9781498710411
9. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (eds.) HCI and Usability for Education and Work. USAB 2008. LNCS, vol. 5298, pp. 63–76. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89350-9_6
10. Vats, V., Nizam, M.B., Liu, M., Wang, Z., Ho, R.: A Survey on Human-AI Teaming with Large Pre-Trained Models. arXiv Preprint arXiv:2403.04931 (2024). https://arxiv.org/abs/2403.04931
11. ISO/IEC 25010:2011. Systems and software engineering – Systems and Software Quality Requirements and Evaluation (SQuaRE). International Organization for Standardization, Geneva (2011)
12. ISO 9241-110:2020. Ergonomics of human-system interaction – Interaction principles. International Organization for Standardization, Geneva (2020)
13. Nielsen, J.: Usability Engineering. Academic Press, San Diego (1993). https://doi.org/10.1016/C2009-0-21512-1

14. Hassenzahl, M.: The Thing and I: Understanding the relationship between user and product. In: Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) Funology. Human-Computer Interaction Series, vol. 3, pp. 31–42. Springer, Dordrecht (2003). https://doi.org/10.1007/1-4020-2967-5_4
15. Theodorou, A., Dignum, V.: Towards ethical and socio-legal governance in AI. Nat. Mach. Intell. **2**, 10–12 (2020). https://doi.org/10.1038/s42256-019-0136-y
16. Amershi, S., et al.: Guidelines for human-AI interaction. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), pp. 1–13. ACM, New York (2019). https://doi.org/10.1145/3290605.3300233
17. Lipton, Z.C.: The mythos of model interpretability. Commun. ACM **61**(10), 36–43 (2018). https://doi.org/10.1145/3233231
18. Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of FAccT 2021, pp. 610–623. ACM, New York (2021). https://doi.org/10.1145/3442188.3445922
19. Mayring, P.: Qualitative Inhaltsanalyse – Grundlagen und Techniken, 13th edn. Beltz, Weinheim (2022)
20. Polit, D.F., Beck, C.T., Owen, S.V.: Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res. Nurs. Health **30**(4), 459–467 (2007). https://doi.org/10.1002/nur.20199
21. Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J.: Evaluating the use of exploratory factor analysis in psychological research. Psychol. Methods **4**(3), 272–299 (1999). https://doi.org/10.1037/1082-989X.4.3.272
22. Kaiser, H.F.: The application of electronic computers to factor analysis. Educ. Psychol. Meas. **20**(1), 141–151 (1960). https://doi.org/10.1177/001316446002000116
23. Cattell, R.B.: The scree test for the number of factors. Multivar. Behav. Res. **1**(2), 245–276 (1966). https://doi.org/10.1207/s15327906mbr0102_10
24. Horn, J.L.: A rationale and test for the number of factors in factor analysis. Psychometrika **30**(2), 179–185 (1965). https://doi.org/10.1007/BF02289447
25. Kaiser, H.F.: An Index of Factorial Simplicity. Psychometrika **39**(1), 31–36 (1974). https://doi.org/10.1007/BF02291575
26. Bartlett, M.S.: Tests of significance in factor analysis. Br. J. Stat. Psychol. **3**(2), 77–85 (1950). https://doi.org/10.1111/j.2044-8317.1950.tb00285.x
27. Brown, T.A.: Confirmatory Factor Analysis for Applied Research, 2nd edn. Guilford Press, New York (2015)

Generational Differences and Perceptions of Artificial Intelligence in Singapore

Chi Tran Mai Nguyen(✉), Chei Sian Lee, Zhang Hao Goh, and Edson C. Tandoc

Nanyang Technological University, Singapore 637718, Singapore
nguyentr004@e.ntu.edu.sg, {leecs,edson}@ntu.edu.sg, zhanghao.goh@tu.edu.sg

Abstract. Each generational cohort - Baby Boomers, Gen X, Gen Y (Millennials), and Gen Z - brings unique experiences and attitudes that shape their expectations of artificial intelligence AI. This study explores these differences and seeks to understand generational attitudes toward AI in the context of Singapore – a prominent technology hub of the world. The study addresses this gap by analyzing survey results from 1,000 adults residing in Singapore based on generations and identifying differences in perception of AI's social impact. Findings indicate primary generational differences in AI usage and perception. The youngest generation Gen Z has the highest level of both usage and knowledge of AI, followed by Millennials. While the majority of the population showed a positive sentiment in perceiving the role of AI in society based on its strong capacity, notable concerns arise for job replacement, fraud, and ethical privacy concerns. AI is expected by the public to transform and advance key sectors, albeit the overall social transformation is under skepticism. Generational perceptions of AI shape attitudes toward adoption and its role in society. Therefore, bridging generational divides in AI perceptions can drive responsible adoption, ensuring AI systems meet the diverse needs of society.

Keywords: artificial intelligence · AI perception · generational differences

1 Introduction

The world has entered a new phase of change as a result of the development of sophisticated information and communication technology, which has created possibilities and solved significant problems in every area of society [1]. With the use of developing technologies, gadgets, and infrastructure, digital transformation is spreading quickly and significantly across a number of organizations, areas, industries, and societies [2]. This change is made possible by information and communication technology (ICT), which manipulates and communicates information processes using computers, electronic sources, Internet tools, and other broadcasting technologies [3]. Among the latest technological advancements, artificial intelligence (AI) is among the most ground-breaking and has a prominent impact on all aspects of the technological and social world. Artificial Intelligence (AI) is the science and engineering of building intelligent computers

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 377–385, 2025.
https://doi.org/10.1007/978-3-031-94171-9_36

that share characteristics with humans [4]. As a result, it entails comparatively intricate computing processes that underlie intelligent machines, including large data, machine learning, and expert-level natural language processing [5].

AI is transforming society, influencing how different generations perceive, interact with, and adopt this technology [6, 7]. Each generational cohort - Baby Boomers, Gen X, Gen Y (Millennials), and Gen Z - brings unique experiences and attitudes that shape their expectations of AI [8, 9]. However, understandings of these experiences and attitudes, along with generational differences in the broader social context, remain limited. This study thus aims to explore these differences and seeks to understand generational attitudes toward AI in the context of Singapore – a prominent technology hub of the world [10]. This study answers the research question: What are the generational differences and perceptions of artificial intelligence among individuals in Singapore?

2 Literature Review

2.1 Public Perception of Artificial Intelligence

Public perception plays a crucial role in AI integration into multiple domains in society. It is defined as the aggregate views of a social group from what its members think of a particular issue or event [11]. Without a thorough understanding of public perceptions, it will be remarkably challenging for technology companies, governance systems, and institutional agencies to universalize the ethical and effective use of AI [12]. Recent studies identified public perceptions toward AI are navigated through a variety of affordances provided by interaction with AI. For instance, the study by Shin [13] explored how causability (the degree to which AI's explanation of a statement in a certain context of usage results in a given degree of causal knowledge with efficacy, efficiency, and satisfaction) and explainability (AI's competency in explaining users' questions with interpretability and understandability) generated better perceptions and trust from AI users. AI agents' fairness also caused an impact on users' impressions through AI-mediated communication, thus enhancing social perceptions and relations of the public towards AI as studied by Li et al. [14].

With a mix of complicated affordances and features of AI, the perception of the public towards AI is also mixed in different contexts. The study by Brauner et al. [15] evaluated the public perspective of social acceptance towards AI as either black or white, but a nuanced mix of both positive and negative expectations and assessments. The counterviews were also acknowledged in the study by Kushwaha et al. [16], in which public perceptions of AI-dependent decision-making varied in risk propensity, especially when interacting with individual aspects such as accountability, self-efficacy, and knowledge of AI. While receiving almost similar interfaces and interaction mechanisms with AI, what causes the notable gap between public groups' usage, knowledge, and perception is worth studying to understand this disparity.

2.2 Generational Differences in AI Public Perception

Age is a crucial factor that accounts for shaping the public perception of AI. Few studies have addressed age and generational differences in fragmented domains rather than

comprehensive social impacts. In education, the study by Chan et al. [7] examined the gap in perception towards AI usage, with Gen Z being more optimistic and open to using AI for better productivity, efficiency, and overall learning experiences than Gen X and Gen Y. Another study by Moravec [8] investigated AI particularly in the context of journalism, with most of the generations (Gen Z, Gen Y, Gen X, and older) leaning towards mixed and negative attitudes toward applying AI in journalistic practices. Barbul and Bojescu [9] focused on the human-AI interaction and the perception of people below and above 30. Their study confirmed the significant generational gap, with Gen Z perceiving AI as more helpful, easy to use, fun, entertaining, pleasant, and willing to use AI-powered services while older generations avoid AI usage when possible [9]. Lee et al.. [17] focused on Gen Z during an academic transition and found the information-seeking behaviors of Gen Z differ from those of their predecessors because of their dependency on digital technologies.

A majority of them profoundly placed importance on Gen Z because of their familiarity and their technological proficiency. Collectively, past research indicates there are generation differences in terms of technological perceptions and usage. Yet, there is still limited understanding regarding specific AI perceptions across generations. Hence, this study aims to address this gap.

3 Methodology

The study addresses the research gap by analyzing survey results from 1,000 participants based on generations and identifying differences in perception of AI's social impact. The research process was under ethical approval from the university's institutional review board to ensure ethical concerns related to the participation of human subjects. All participants were above 21 years of age by the time of data collection and received financial compensation upon completion.

Besides demographic information, the survey asked participants about usage, knowledge, and perceptions of AI tools which were adapted from Bach et al. [18] and Kennedy et al. [19]. AI interaction frequency measured how often participants interacted with AI on a five-point Likert scale. AI perceived knowledge measured the level of adequacy participants evaluated themselves in the knowledge of AI for everyday use on a five-point Likert scale. AI information-seeking platforms measured prior usage of any AI online or offline platforms participants used to search about AI (as a dichotomous variable) and the level of trust in such platforms (on a three-point Likert scale). AI future prediction measured the level of positivity participants had on the future impact of AI on various aspects of life, including employment, healthcare, privacy, and society on a five-point Likert scale. The questionnaire also asked participants about keywords associated with their perceptions of AI.

Regarding the study sample, the total responses used for the study were 1,000 adults above 21 years old. The sample was approximately balanced in gender (52.9% male, 47.1% female). Participants were all Singaporean residents, either Singaporean citizens (89.3%), permanent residents (9.7%), or non-Singaporeans (1%). Based on participants' age ($M = 45.94$, $SD = 13.53$), generations were categorized as Gen Z from 21–27 years old (7.2%), Gen Y from 28–43 years old (38.9%), Gen X from 44–59 years

old (36.1%), and Baby Boomers and above from 60 + years old (17.8%). The sample's demographic represented the actual Singaporean total population. For data analysis, the study employed descriptive analysis using IBM SPSS 28 software and visualized the data using Excel and Voyant.

4 Findings and Discussions

The analysis results reflect prominent gaps between generations in AI usage, knowledge, and perceptions in the Singaporean sampled population. Regarding interaction frequency as indicated in Fig. 1, Gen Z shows an outstandingly higher level of usage ($M = 3.17$, $SD = 1.20$), followed by Gen Y ($M = 2.30$, $SD = 1.30$). Interestingly, Gen X has the lowest perceived interaction frequency ($M = 1.99$, $SD = 1.28$), lower than the older generation – Baby Boomers & older ($M = 2.21$, $SD = 1.37$). Younger generations including Gen Z and Gen Y show a higher tendency to use and interact with AI more frequently. This result aligns with previous studies [7, 9], showing that the young population is more tech-savvy and adaptive to new technologies.

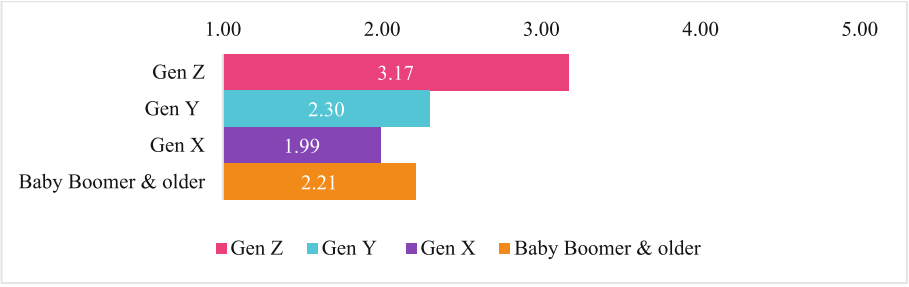


Fig. 1. Mean differences of AI interaction frequency by generation.

In terms of perceived knowledge of AI, the population shows a skewness towards the younger population, as illustrated in Fig. 2. Baby Boomers and older groups are least confident in their knowledge of AI ($M = 2.47$, $SD = 1.20$), while Gen Z is the most confident group ($M = 3.39$, $SD = 0.99$), followed by Gen Y ($M = 2.92$, $SD = 1.11$) and Gen X ($M = 2.68$, $SD = 1.08$). Younger demographics have a higher level of self-efficacy than their older counterparts, which also aligns with previous findings in evaluating the perceived level of understanding of AI [7]. This is presumably contributed by the quick adaptation and applications by the young groups to the new technologies, in which AI is part of the latest universalized technological advancements [9].

The textual visualization in Fig. 3 offers a more nuanced viewpoint on the differences in sentiments toward AI between generations, highlighting the varying attitudes toward AI. Gen Z and Gen Y are remarkably more optimistic and positive towards AI, perceiving it as a powerful tool for meaningful purposes, with keywords like "good", "cool", and "smart" having the highest frequency in these age groups. In contrast, older generations (Gen X, Baby Boomers, and older) exhibit a more cautious stance towards AI, especially with the high number of recorded instances of the keyword "worrying" (or

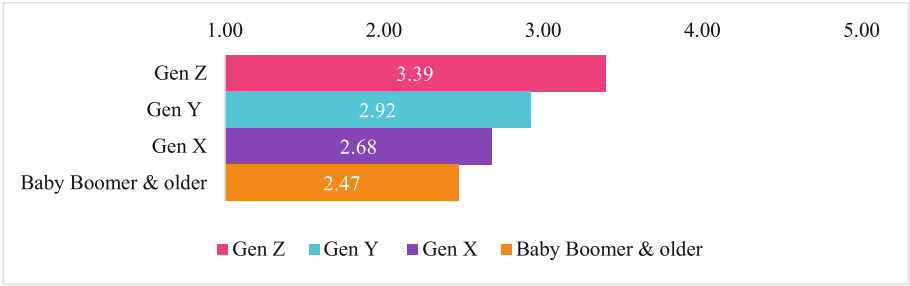


Fig. 2. Mean differences of AI perceived knowledge by generation.

"worry", "worried"). These generations balance curiosity with concern about the ethical and societal implications of AI. The result is similar to previous studies [7, 9], indicating a more open-minded and progressive perception of AI from the younger groups compared to their senior counterparts. This nuanced divergence underscores the generational perspectives as AI continues to evolve and impact various aspects of life.

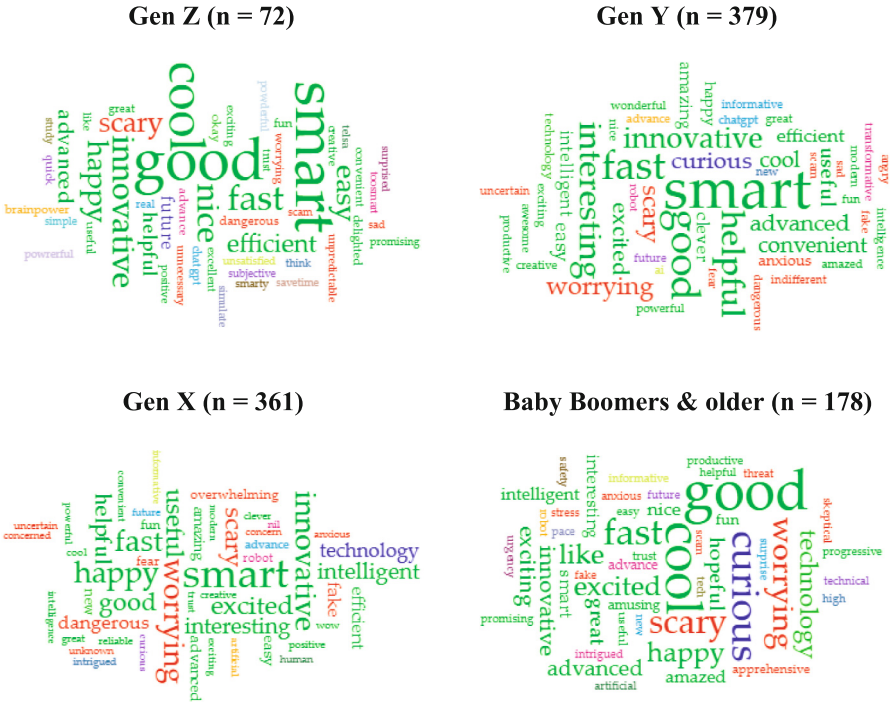


Fig. 3. Word cloud of adjectives describing feelings or emotions when hearing the phrase Artificial Intelligence (AI) by generation (n = 1,000). Note. 50 terms most mentioned; size reflects frequency; color indicates sentiments.

The perception of different generations on the future impact of AI on society also varies across domains, as depicted in Fig. 4. Firstly, in terms of employment, younger generations including Gen Z ($M = 2.55$, $SD = 0.68$) and Gen Y ($M = 2.43$, $SD = 0.75$) show a more notably negative attitude towards AI compared to the older generations with Gen X ($M = 2.57$, $SD = 0.67$), Baby Boomers and older ($M = 2.68$, $SD = 0.62$), although all generations show a general positive prediction towards the issue. Millennials are the most negative towards this issue, while Baby Boomers and older people are the most positive. This finding contrasted with some previous findings on the acceptance and usage of AI for professional purposes [7]. However, this can be understood in how each generation perceives AI with job security. The older generations can expect AI to be more efficient and helpful in assisting human workers to enhance productivity, as well as foster technological advancement on a macro scale to create more jobs for more people. Younger generations with less work experience can be more concerned with AI replacing their work that doesn't require extensive experience, especially in management roles.

Secondly, regarding healthcare, there are no prominent differences between generations, with all age groups predicting a fairly positive impact of AI on healthcare. While Gen X ($M = 2.34$, $SD = 0.73$) is the most positive, followed by Gen Y ($M = 2.26$, $SD = 0.75$), Gen Z ($M = 2.23$, $SD = 0.72$), Baby Boomers and older ($M = 2.23$, $SD = 0.72$) have the least positive predictions on this concern. The overall optimism resonates with the research by Gao et al. [20] on public perception towards medical AI, in which both indicate positive reactions but are not too highly preferred. This can be attributed to the technology still under development, which is yet to be fully universalized along with controversies on effectiveness and ethics [20]. Thirdly, in terms of the prediction of AI in privacy matters, the differences between each generation are more notable. Younger generations express an explicit concern towards privacy, with Gen Z being the most skeptical Gen Z ($M = 2.44$, $SD = 0.78$), followed closely by Gen Y ($M = 2.53$, $SD = 0.71$). Senior generations show a more positive prediction of privacy matters, with Gen X having the most favored evaluation Gen X ($M = 2.65$, $SD = 0.61$) along with Baby Boomers and older ($M = 2.58$, $SD = 0.63$). The results reflect the general divide between generations with different levels of digital and AI literacy. Concerns related to privacy on various contemporary technologies have been notable among the younger generations, especially Gen Z who have been exposed to such advancement at an early age and thus have a higher overall digital literacy.

Finally evaluating the impact of AI on society in general, the generational divide is fairly significant and worth spending the attention for. Only Gen Z predicts that AI will facilitate a better future for society ($M = 2.05$, $SD = 0.82$), even though the mean evaluated score exceeds the average benchmark by a relatively small margin. Gen Y ($M = 1.90$, $SD = 0.78$), Gen X ($M = 1.75$, $SD = 0.75$), Baby Boomers, and older ($M = 1.78$, $SD = 0.77$) all perceive that AI will affect society for the worse. This result aligns with the current viewpoint in research regarding the divide in generational perceptions of AI with a more favored endorsement from Gen Z compared to the senior groups [7, 9].

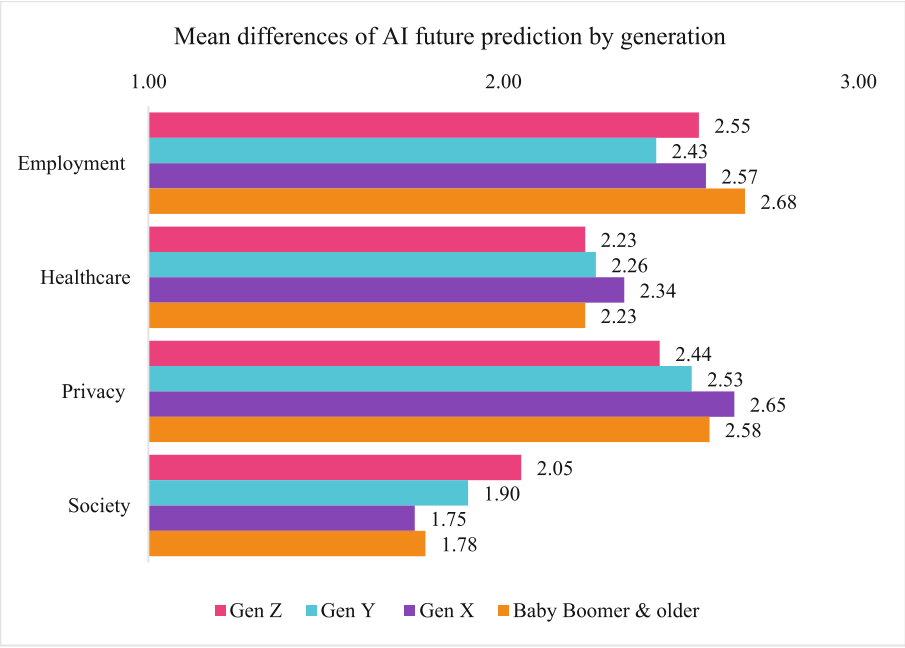


Fig. 4. Mean differences of AI future prediction by generation. Note. Higher values indicate more positive predictions.

5 Conclusion

The study explores the public perception of the AI impact on society, with a particular focus on the context of Singapore as an AI-driven nation [10]. Overall, findings indicate primary generational differences in AI usage and perception. The youngest generation Gen Z has the highest level of both usage and knowledge of AI, followed by Millennials. Younger generations (Gen Z and Millennials) also possess a more proactive approach towards AI with a higher tendency to look up AI information. While the majority of the population showed a positive sentiment in perceiving the role of AI in society based on its strong capacity, notable concerns arise for job replacement, fraud, and ethical concerns for privacy. AI is expected by the public to transform and advance key sectors, albeit the overall social transformation is under skepticism.

5.1 Theoretical & Practical Implications

Regarding theoretical contributions, this study addresses the research gap in understanding public usage, knowledge, and perceptions of AI's social impact on a national representative sample. More importantly, the distinct characteristics between generations accounted for the difference between different demographics, thus offering an insightful perspective for future studies to acknowledge this distinction in designing research on AI public perceptions.

Findings on generational perceptions of AI shape attitudes toward adoption and its role in society also have practical contributions. Specifically, the findings can inform policies for schools and workplaces, which help bridge generational divides in AI perceptions. This will ultimately drive responsible adoption, ensuring AI systems meet the diverse needs of society.

5.2 Limitations & Future Research

The study encounters the following limitations. First, the cross-sectional study design limits the ability to explore the explicit transformation of public perception toward the positive or negative side. Future research can address this limitation through a longitudinal research design with multiple waves of data collection and iterative analysis for a more comprehensive viewpoint of the transformation. Second, a causal explanation of such perception of AI and its generational differences has not been addressed. Studies that can examine the determinants of AI public perception, considering various contextual and technological factors will provide a more comprehensive viewpoint on this researched area.

Acknowledgments. This study was funded by ai.sg (grant number AISG3-GV-2021–003).

Disclosure of Interests.. There are no competing interests among all the authors.

References

1. Sarirete, A., Balfagih, Z., Brahimi, T., Lytras, M.D., Visvizi, A.: Artificial intelligence and machine learning research: towards digital transformation at a global scale. J. Ambient. Intell. Humaniz. Comput. **13**(7), 3319–3321 (2022)
2. Zhu, X., Ge, S., Wang, N.: Digital transformation: a systematic literature review. Comput. Ind. Eng. **162**, 107774 (2021)
3. Lwoga, E.T., Sangeda, R.Z.: ICTs and development in developing countries: a systematic review of reviews. Electron. J. Inf. Syst. Dev. Countries **85**(1), e12060 (2019)
4. Monostori, L., Artificial Intelligence, in CIRP Encyclopedia of Production Engineering, Laperrière, L., Reinhart, G., Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 47–50
5. Mondal, B.: Artificial intelligence: state of the art. In: Balas, V.E., Kumar, R., Srivastava, R. (eds.) Recent Trends and Advances in Artificial Intelligence and Internet of Things, pp. 389–425. Springer, Switzerland (2020)
6. Harper, R.H.: The role of HCI in the age of AI. Int. J. Human-Comput. Interact. **35**(15), 1331–1344 (2019)
7. Chan, C.K.Y., Lee, K.K.: The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? Smart Learn. Environ. **10**, 60 (2023)
8. Moravec, V., Hynek, N., Skare, M., Gavurova, B., Kubak, M.: Human or machine? The perception of artificial intelligence in journalism, its socio-economic conditions, and technological developments toward the digital future. Technol. Forecast. Soc. Chang. **200**, 123162 (2024)

9. Barbul, M., Bojescu, I. Generations' perception towards the interaction with AI. In: Pamfilie, R., Dinu, V., Vasiliu, C., Pleșea, D., Tăchiciu, L. edn. Proceedings of the BASIQ International Conference, pp. 539–546. Bucharest: ASE (2023)
10. Pande, D., Taeihagh, A.: A governance perspective on user acceptance of autonomous systems in Singapore. Technol. Soc. **77**, 102580 (2024)
11. Dowler, E., Green, J., Bauer, M., Gasperoni, G.: Assessing public perception: Issues and methods. In: Dora, C. (ed.) Health hazard and public debate: lessons for risk communication from BSE/CJD saga, pp. 40–60. World Health Organization, Geneva (2006)
12. Fast, E., Horvitz, E.: Long-term trends in the public perception of artificial intelligence. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 963–969 (2017)
13. Shin, D.: The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. Int. J. Hum. Comput. Stud. **146**, 102551 (2021)
14. Li, J., Chu, Y., Xu, J.: Impression transference from AI to human: the impact of AI's fairness on interpersonal perception in AI-Mediated communication. Int. J. Hum. Comput. Stud. **179**, 103119 (2023)
15. Brauner, P., Hick, A., Philipsen, R., Ziefle, M.: What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI. Front. Comput. Sci. **5**, 1113903 (2023)
16. Kushwaha, A.K., Pharswan, R., Kumar, P., Kar, A.K.: How do users feel when they use artificial intelligence for decision making? A framework for assessing users' perception. Inf. Syst. Front. **25**(3), 1241–1260 (2023)
17. Ho, M.-T., Mantello, P., Ghotbi, N., Nguyen, M.-H., Nguyen, H.-K.T., Vuong, Q.-H.: Rethinking technological acceptance in the age of emotional AI: surveying Gen Z (Zoomer) attitudes toward non-conscious data collection. Technol. Soc. **70**, 102011 (2022)
18. Bach, T.A., Khan, A., Hallock, H., Beltrão, G., Sousa, S.: A systematic literature review of user trust in AI-enabled systems: An HCI perspective. Int. J. Hum.-Comput. Interact. **40**(5), 1251–1266 (2024)
19. Kennedy, B., Tyson, A., Saks, E.: Public awareness of artificial intelligence in everyday activities, in American trends panel wave 119. Pew Research Center (2023)
20. Gao, S., He, L., Chen, Y., Li, D., Lai, K.: Public perception of artificial intelligence in medical care: content analysis of social media. J. Med. Internet Res. **22**(7), e16649 (2020)

Towards a Design Taxonomy of AI Products and Services

Simon Norris(✉), Andrea Resmini, Jesper Lund, and Susanne Lindberg

School of Information Technology, Halmstad University, Kristian IV: S Vag 3, Halmstad, Sweden
simon.john.norris@hh.se

Abstract. The recent explosion in artificial intelligence and its integration into more and more domains beyond traditional ones such as health care, insurance, and defence show no signs of slowing down. Generative AI is becoming part and parcel of an increasing number of designerly activities as either a tool—for example in the form of the Large Language Generative Models that enable the creation of text, images, videos, or mock-ups via chatbot-like interfaces—or a material—for example when creating the interfaces that others will use to interact with such models. Rapid adoption and low AI literacy among designers' limits awareness of limitations, challenges, opportunities, and consequences deriving from the use of AI-enabled products and services. The paper introduces the concept of Design Literacy for AI as distinct from AI Literacy and Design Literacy, and proposes a design-centred taxonomy of Generative AI products and services. The taxonomy is operationalised into a proof-of-concept web service using faceted classification to help designers better understand AI as both a tool and a material, further investigate the concept of Design Literacy for AI, and contribute to reducing the gap between AI development and design practices.

Keywords: Design Literacy for AI · AI · Taxonomy · Faceted Classification

1 Introduction

The emergence of generative AI (GAI) as a new technology has "extended traditional, data-driven AI tasks such as predictions, classifications, or recommendations toward the generation of unique, realistic, and creative content" (Banh & Strobel, 2023, p. 1). GAI is profoundly different from previous flavours of AI and can be considered "an umbrella term for models that create novel output, either independently (randomly) or based on a prompt supplied by the user" (Kneusel, 2024, p. 97). Prior to the rise of GAI, predictive AI (PAI) was the dominant contemporary form of AI, relying on machine learning algorithms that analyse historical data to identify patterns and forecast future outcomes (Siegel, 2013; Narayanan & Kapoor, 2024).

PAI concerns itself with specialised tasks such as market trend prediction, fraud detection, and customer behaviour analysis, whereas GAI is a content creation technology that learns from existing data being integrated into a growing range of products and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 386–396, 2025.
https://doi.org/10.1007/978-3-031-94171-9_37

services. Narayanan and Kapoor (2024) outline that whilst PAI excels at pattern identification and decision-making support, GAI demonstrates impressive creative capabilities that can nonetheless result in misinformation.

A key distinction between these forms of AI is their capacity for generalisation. Goetz et al. (2024) define generalisation as "the ability of AI systems to apply and/or extrapolate their knowledge to new data which might differ from the original training data". PAI development has been aimed at task-specific, narrow use cases, whereas the development of the foundation models that power GAI products and services explicitly aims at their generalised application across diverse domains, contexts, and use cases. Such generalisable foundation models underpin a growing number of widely-used GAI products and services, including the Generative Pre-Trained Transformer (GTP) series by OpenAI and LLaMA (Large Language Model Meta AI) by Meta that apply Natural Language Processing (NLP). Understanding how these models work, what their capability and limitations are once they are moulded into products or services, represents a complex space for designers to navigate, where different types of knowledge and practice overlap. There is technical knowledge associated with designing AI, the knowledge to integrate the models into a product or service, and the knowledge necessary to proficiently use such products and services effectively from a human perspective.

The fact that a Human-Centred Design (HCD) perspective is necessary for creating serviceable and useful products and services is not disputed (Chen et al., 2022). Still, the nature of AI development creates a grey zone in which many design interventions happen at a superficial level—that of the final user interface—when production choices have already shaped design features (labelling, tone of voice), or not at all. Fundamental questions related to creating (G)AI systems that align with human values and needs, bridging the gap between technical capabilities and societal requirements, remain to be clarified. This grey zone is the friction point where the two very different perspectives of AI development and HCD practices on epistemology, processes, organisation, and human-centredness meet (Table 1).

In terms of *epistemology*, AI development combines deductive and inductive technical approaches, emphasising precision and computational accuracy. The goal is accuracy, based on performance and quantifiable outcomes. Design practices are abductive and focus on the 'good fit' between artifact and context (Alexander, 1964): knowledge is generated through the exploration of possibilities. The AI development *process* is lengthy, linear, highly technical, and involves precise steps: data acquisition, data conditioning, algorithm selection, model training, testing, optimization. HCD involvement occurs either very late in the process (Fischer et al., 2021) or not at all (De Silva & Alahakoon, 2022). Design is driven by user research and centres on understanding people's needs and values. It works through shorter, faster, iterative processes. In terms of *organisation*, AI development requires highly technical set-ups, uses specialized language, and exerts disciplinary control. Design activities have lower technical requirements, use generic language, and accept ambiguity and ill-defined problems. AI development concerns itself with actual use cases and the ethical and societal consequences of an AI model very late in the process, typically after testing or deployment. Guidelines applied to shape requirements focus on technical outcomes to be measured quantitatively. Design practices are people-focused: use and ethical and societal considerations are important

aspects of the contextual creation of the designed artifact. Guidelines are qualitative in nature and consider subjective elements such as mood, feelings, experience (Benyon, 2019).

Table 1. Differences between AI Development Perspective and Design Practices Perspective.

	AI Development Perspective	Design Practices Perspective
Epistemology	Deductive, Inductive Objective Quantitative Accuracy, Performance	Abductive Subjective, Context-specific Qualitative Fit
Process	Longer development lifecycle Driven by data	Faster iterative lifecycle Driven by user research
Organisation	Technical language Disciplinary control Centres on the model	Conceptual language Interdisciplinary collaboration Centres on the person's perspective
Human-centredness	Considered later in the process Quantitative guidelines are applied	Considered throughout the process Qualitative guidelines are applied

Both perspectives would benefit from a better understanding of the other's approach and mindset, but design faces the most immediate disruption. GAI is becoming commonplace in design work, both as an end-product and as a support tool in the creative process. Such rapid adoption has made it difficult to keep up with the speed of change, reducing awareness of limitations, challenges, opportunities, and consequences (Long & Magerko, 2020). To provide designers with a frame for creating 'good fits' when working with and for GAI, an increase in literacy is essential (Lutnæs, 2021; Kiernan et al., 2022). This paper investigates how Design Literacy for AI can be supported by a proposed design-centred taxonomy of GAI products and services. This taxonomy is operationalised into a proof-of-concept web service meant to help designers better understand AI as a material and thereby address the literacy gap between AI development and HCD practices.

2 Literacy, Digital Literacy, and AI Literacy

Literacy is traditionally defined as the ability of "reading and writing effectively in a variety of contexts" (Pilgrim & Martinez, 2013, p. 60) and is often used as a shortcut for literacy, articulacy, numeracy, and those additional necessary competences that are essential for navigating the world (Balchin, 1972). More recently, the importance of engaging with digital information to live and work in a modern society have led to the introduction of the concept of 'digital literacy' (Gilster, 1997; Bawden, 2008). Digital literacy is more than just an extension of the literacy concept to a different context: it is a developmental skill that includes multiple literacies addressing socio-emotional aspects and cultural perspectives related to critical thinking (Ng, 2012; Bowden, 2008), and the

ability to transgress cultural, institutional and territorial borders (Eshet-Alkalai, 2004). While numerous definitions of digital literacy exist, they all centre on those "skills, knowledge and attitudes required to access digital information effectively, efficiently, and ethically" (Julien, 2018, p. 2443).

The increasing integration of AI into all sorts of digital systems (Riedl, 2019) has led to the conceptualisation of 'AI literacy'. Long and Magerko (2020, p. 598) define it as "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace". Importantly, AI literacy concerns itself not only with how AI operates, but also with how it is designed, what limits and underlying biases it has, and the broader social implications thereof (Long & Magerko, 2020). AI literacy includes a multiplicity of sub-literacies that cover technical, cognitive, and ethical factors (Ng et al., 2021) and that go beyond the mere technical to include critical thinking considerations, making AI a topic of education and application in fields and disciplines beyond computer science. One of these fields is design, where integration of AI literacy is essential (Kiernan et al., 2022) but poses questions in respect to the field's own specific "design literacy" and the ways it has been variously framed.

Pacione (2010) intends being "design literate" as possessing skills that "are well within the full range of everyone's cognitive and kinaesthetic capabilities and serve our everyday needs", such as "inquiry, evaluation, ideation, sketching, and prototyping", and considering "literate" in opposition to "expert" or "professional". Design literacy is in this sense a foundational type of literacy, not just for designers, but for everyone "whose job it is to imagine something that does not yet exist and then plot the path from imagination to existence" (Pacione, ibid.). Lutnæs (2021, pp. 2–3) agrees, but considers how a sustainable concept of design literacy also needs to include "critical scrutiny of how a newly proposed innovation supports a transition towards a more sustainable future" not to remain simply a "driver of economic competitiveness". This requires us to introduce "critical reflection as a baseline to any design process" and "asking basic ethical questions of how a new idea makes the world better", an aspect of design literacy that is crucial for ensuring that AI-driven innovation remains ethical, people-centred, and sustainable.

The focus on critical scrutiny and societal impact also raises the importances on considering the unintended consequences that can result from AI-enablement (Broussard, 2018), something Human-Centred AI (HCAI), an approach to AI development that focuses on the safety and trustworthiness of AI from the perspective of human needs and values (Shneiderman, 2022), is set to address. Still, HCAI points to a combination of AI literacy and Design literacy that does not yet exist, even as it is cast as vital for critically engaging with and shaping more humane AI applications.

3 AI Literacy Among Designers

Recent analysis of industry practices (Hartikainen et al., 2022) confirms that AI development follows a traditional software development lifecycle and that AI literacy is low among designers (Long & Magerko, 2020). HCD practices, with their focus on the human element, and computer and data science, with its emphasis on predictive performance

based on data selection, data wrangling, and model training, do not share a language, approaches, or outlooks, and this results in misalignments related to timelines, language, focus, and organisation.

Timeline mismatches happen primarily because of AI's longer data/model cycles, which contrast to the shorter turnaround of design activities centred on prototyping, testing, and rapid iterations (Benyon, 2019). Differences in language and vocabulary are a common issue (Yang et al., 2020). AI development focuses on metrics related to its pipeline and accuracy of output; design practices concern themselves with broader, harder to define societal and ethical issues (Buchanan, 1992; Resmini, 2012; Lindenfalk & Resmini, 2022). AI development teams and design teams are organised differently and kept separated, undermining the interdisciplinarity that is intrinsic to the concept of AI literacy (Yang et al., 2020). This often results in the relegation of design concerns to the later stages of the development process where focus rests entirely on user interfaces (UI) and surface elements rather than on the model conception (Fischer et al., 2021). Ideally, from a design perspective, user research should inform the data collection and model selection stages, so that HCAI characteristics such as transparency and interpretability are addressed early on (Benyon, 2019; Shneiderman, 2022). Current organization models mean that fundamental issues of explainability, trust, and user experience may be entirely left out of the development lifecycle.

These misalignments are made more pervasive by the fact that designers still struggle to grasp AI (Yang et al., 2020). Low AI literacy and a lack of familiarity produce uncertainty about how AI impacts design processes (Windl et al., 2022) and how design could or should impact AI processes (Crawford, 2021), resulting in immature AI-centred design practices (Hartikainen et al., 2023).

Improving such a situation requires designers to become familiar with current types and flavours of AI, their capabilities, applications areas, and algorithms, so that they can conceptualise AI in terms of design work and of how design methodologies can be integrated into, support, and improve existing AI development lifecycles and result in 'good fit' products and services. The issue is then whether all designers need is to improve their AI literacy.

Current AI development processes involve three different types of 'design': the first type covers the development of AI, and is concerned with the creation of a model or AI product (Fischer, 2021; De Silva & Alahakoon, 2022). For example, the creation of GPT; the second type concerns itself with how a newly created AI model can be presented so that people can interact with it. For example, designing an interface for GTP and creating ChatGPT; the third type focuses on how people use the AI products and services to improve, shape, and influence their own work. For example, using ChatGPT to produce a brief. These three types are respectively defined in this paper as "Designing AI' (DAI), 'Designing for AI' (DfAI), and 'Designing with AI' (DwAI).

As Table 2 shows, these three activities not only differ in focus and in what they set out to produce, but require different levels of AI literacy and Design literacy, and employ different types of knowledge in action. In terms of focus, DAI is currently a software engineering process entirely centred on data, algorithm selection, model training, and optimisation. DfAI is concerned with the UI, the user's experience, and explainability. DwAI is concerned with integrating AI products in design processes for content and

Table 2. Design Literacy for AI.

	Designing AI (DAI)	Designing for AI (DfAI)	Designing with AI (DwAI)
Focus	Data, algorithms, models	An AI-enabled product or service	Use of an AI-enabled product or service to make something else
Level of AI Literacy	Expert	Competent to expert	Novice to competent
Level of Design Literacy	Novice to competent	Expert	Competent to expert
Knowledge Types	Deductive, inductive	Abductive	Pragmatic

code creation, prototyping, and ideation. In terms of knowledge, DAI works deductively and inductively with data-driven models, focusing on performance to measure accuracy and precision. DfAI emphasises abductive knowledge to shape 'good fit' UIs and user experiences in context. DwAI emphasises pragmatic knowledge on how to apply AI-enabled products or services in creative or problem-solving activities. Literacy, both AI literacy and Design literacy, is here coded using the novice, competent, and expert stages from Dreyfus Skill Model (Rousse & Dreyfus, 2021): being a novice implies knowledge of general rules abstracted from engagement in a domain; being competent means being responsive to more dimensions of a situation and choosing a perspective; being an expert means 'seeing what to do' and closing the gap between intuitive perspective and action.

4 Design Literacy for AI

Table 2 shows how the requirement for AI literacy is highest when designing AI, it decreases when designing for AI, and is at its lowest when designing with AI. This reflects a common-sense understanding of the production process, where specialised technical knowledge is necessary to develop a marketable artifact (DAI, DfAI) but not to use it (DwAI). This is not the case for Design literacy, which can be said to be novice to competent for DAI, if there at all, but then is at expert level in DfAI activities and remains competent to expert in DwAI. This reflects two very different literacy needs for designers: being fully conversant with AI so they can create the best 'fit' for others (DfAI); being sufficiently conversant with AI to avoid 'bad fits' and unintended consequences when using it themselves to design other artifacts (DwAI).

We collectively call these two types of literacy 'Design literacy for AI' (DL4AI), which we define as the ability to understand, apply, and critically engage with design principles, methods, and practices across DAI, DfAI, and DwAI activities.

Exploring DL4AI implies exploring what it means to be fully conversant in AI from a perspective epistemologically distant from that of AI development, working through different processes organised differently, to answer different concerns. Following well-established design practices (Design Council, n.d.), exploration of the concept began with a discovery phase that led to the idea of a discovery tool that could be used to

quickly investigate, compare, and sense-make. As understanding begins with the ability to name, compare, and differentiate, a taxonomy of AI products and services in the GAI space was identified as the necessary starting point.

5 A Taxonomy for Design Literacy for AI

A number of recent papers have introduced taxonomies that describe AI from the perspective of software development (Barredo Arrieta et al., 2020; EIT, 2021; De Silva & Alahakoon, 2022; Lu et al., 2024). These were initially used as blueprints for the DL4AI taxonomy presented in this paper specifically focusing on a sample database of GAI products and services, in response to the current attention these are receiving and because of their disruptive relevance in design practices. The sample data for this paper came from a desk survey of existing AI products and services and from a review of current literature on AI, DfAI and DwAI. Initial iterations of the taxonomy made it immediately clear that the gap between these technology-centred efforts and the mindset and practice of designers prevented them from being effective support for DL4AI. Early categories such as 'system architecture' or 'learning paradigm' did not provide knowledge insights in the abductive and pragmatic space of DfAI and DwAI.

On one hand, this confirmed the gap between technical perspectives of AI and more human-centred perspectives. A data scientist may think of AI in terms of the learning paradigm or base algorithm it uses, but designers may want to understand how they would need to interact with it, via a keyboard, voice, or other input types. This required organising the AI domain in radically different ways from those of software engineering literature, focusing on characteristics such as interaction, use, or price.

On the other hand, the sheer breadth of possible choices clearly pointed to the need of conceptualising a web service that would support structured free exploration of what would be a very large and sprawling database. Designers may also have complex, pragmatic questions such as "what image generators produce photorealistic images", "which AI products were trained on open data", or "which text-to-image models are free with and without an account". This suggested that the information architecture supporting the web service should organise the taxonomy into a faceted classification.

Figure 1 shows the current seven primary facets used by the web service. The "Risk Level" facet follows the risk-based classification introduced by the EU AI Act (European Commission, 2024).

Faceted information architectures have the remarkable property of allowing continuous filtering and de-filtering to zoom in or out of any result set (Quintarelli et al., 2007). Faceted or colon classification (Ranganathan, 1933) can combine multiple categories, the facets, into dynamic sectioning principles to order items in one or more domains and offer different views of the content (Benyon, 2019). Facets work as filtering sieves, allowing a "focus / magnify" approach and preventing empty result sets (Resmini & Rosati, 2010; Resmini & Rosati, 2011).

Importantly, facets can be added to handle items that introduce new characteristics. This makes them best suited for a web service meant to sense-make in a rapidly evolving space where language and definitions change daily, making consolidated taxonomies problematic.

Product	Model	Capabilities	Pricing	Openness	Risk Level	Interaction
ChatGPT	Animation & Motion Capture Models	Accessibility Features	Annual Subscription	API vs. Full Access	Minimal risk	API Access
Claude	Audio-Based Generation	Assisted Writing	Enterprise Plan	Commercial Use	Acceptable risk	Canvas Mode
DALL-E	Conversational	Chat-based UX	Free Trial	Data Retention Policies	High risk	Command Line
DeepSeek R1	Creative Writing Models	Decision Making	Free With Account	Data Usage Transparency	Unacceptable risk	Contextual Memory
ElevenLabs	DeepFake Models	Design Elements	Free Without Account	Modification Rights		Custom UI Elements
Firefly	Image-Based Generation	Generative Art	Monthly Subscription	Open Community Support		File Upload/Download
Gemini	Image-Editing & Enhancement	Image Editing	Multiuser Plan	Open Source		Image Generation
Jasper	Image-Synthesis Models	Prototyping	One-Time Purchase	Proprietary		Multimodal Input
LLaMA	Music Composition Models	Storyboarding	Pay-Per-Use	Training Data Transparency		Prompting
Midjourney	Sound Effect Generation Models	Supported Research		Usage Rights		Text Editing Controls
Runway AI	Speech Synthesis & Cloning Models	UI Sound Effects		Weight Accessibility		Voice Interaction
Stable Diffusion	Text Completion	Voice				
	Text-Based Generation					
	Video-Based Generation					

Fig. 1. The HCD-centred facets of the taxonomy, with sample data.

6 The Web Service

The taxonomy is currently conceptualised as a web-based openly accessible service that allows visitors to explore a small sample of GAI products and services, exposing their relationships through the logic of a filtering interface based on HCD-centred facets that allow visitors to zoom into the data. Figure 2 shows a mock-up of the web service interface after a person selects values for 'Model' and 'Pricing' to see which free 'text-to-image' generators exist. A conversational rendition of the query is visible top-left. Three generators are shown at the top, and 'Stable Diffusion' is selected, so the main body shows more detailed information on it. From here, one could zoom in using any choice still available, for example 'Team Plan' in the 'Pricing' facet would further restrict the result set to only show 'Stable Diffusion'. Or one could zoom out, deselecting the 'Pricing' checkboxes to see a list of all 'text-to-image' generators, which will also include, for example, Dall-E. Unavailable choices are greyed out.

7 Conclusions and Future Work

Designers struggle with AI literacy, which creates uncertainty about how AI impacts design processes and how design could or should impact AI processes (Crawford, 2021). Addressing this literacy gap is crucial to ensuring that AI-empowered products and services are designed to be safe, reliable and trustworthy (Shneiderman, 2022).

DL4AI is a conceptual framework that describes the ability to understand, apply, and critically engage with AI from a design perspective, and that emphasises the designer's role in shaping AI interactions and experiences. We categorise AI-related design as DAI, DfAI, and DwAI, each requiring different levels of AI and design literacy. DAI focuses on the model: from a design perspective, it suffers from the lack of human-centred considerations. DfAI focuses on UIs and the experiences of people interacting with AI-enabled systems: designers crucially have to make AI understandable, interpretable, and usable. DwAI focuses on using AI products in design work: it is about proficiency in using AI as augmentation in design practice.

We propose a HCD-centred taxonomy ordered by products, AI models, capabilities, interactions, risks, price, and openness. This taxonomy shifts the focus from technical

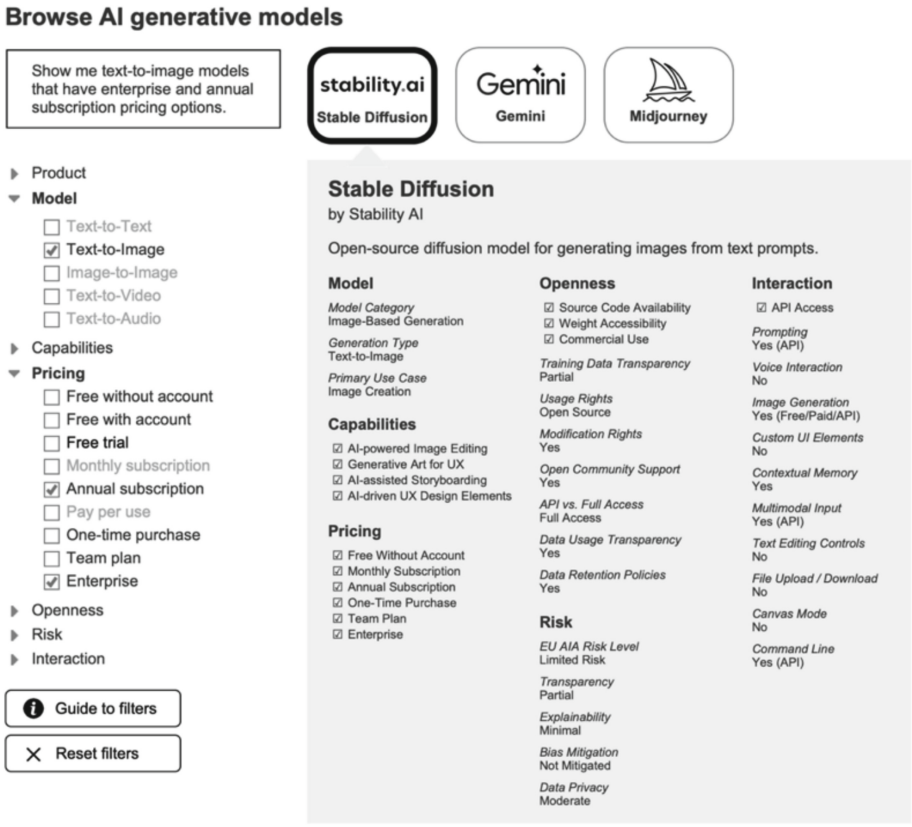


Fig. 2. A mock-up of the web interface for querying the GAI taxonomy.

classifications to practical, people-oriented criteria that suit designers who have to make informed decisions about AI integration. We also illustrate a web-based service proof of concept that provides designers with an interactive way to explore the taxonomy, enabling them to make sense of a complex evolving landscape by adjusting their focus based on their specific needs.

Future work involves extending the taxonomy to cover a wider array of AI products and services outside of the GAI pool, keeping the database up to date, testing to ascertain whether more or different facets are needed, and if additional information should be included to further support DL4AI. In the longer term, expand the concept of DL4AI, and assess whether DL4AI could be used to bridge the perspective and process gap between DAI and DfAI / DwAI.

Disclosure of Interests. The authors have no relevant competing interests to declare.

References

1. Alexander, C.: Notes on the Synthesis of Form. Harvard University Press, Cambridge (1964)

2. Balchin, W.G.V.: Graphicacy. Geography **57**(3), 185–195 (1972). http://www.jstor.org/stable/40567805
3. Banh, L., Strobel, G.: Generative artificial intelligence. Electron. Markets **33**(63) (2023). https://doi.org/10.1007/s12525-023-00680-1
4. Barredo Arrieta, A., Diaz-Rodriguez, N., Del Ser, J., et al.: Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fus. **58**, 82–115 (2020). https://doi.org/10.1016/j.inffus.2019.12.012
5. Bawden, D.: Origins and concepts of digital literacy. In: Lankshear, C., Knobel, M., (eds.) Digital Literacies: Concepts, Policies and Practices, pp. 178–32. Peter Lang (2008). https://doi.org/10.1163/9789004265564_002
6. Benyon, D.: Designing User Experience: A guide to HCI, UX, and interaction design (4th ed). Pearson (2019)
7. Broussard, M.: Artificial unintelligence: how computers misunderstand the world. The MIT Press (2018)
8. Buchanan, R.: Wicked problems in design thinking. Des. Issues **8**(2), 5–21 (1992). https://doi.org/10.2307/1511637
9. Chen, Y., Clayton, E.W., Novak, L.L., et al.: Human-centered design to address biases in artificial intelligence. J. Med. Internet Res. (2022). https://doi.org/10.2196/43251
10. Crawford, K.: Atlas of AI. Yale University Press, New Haven (2021)
11. De Silva, D., Alahakoon, D.: An artificial intelligence life cycle: From conception to production. Patterns **3**(6) (2022). https://doi.org/10.1016/j.patter.2022.100489
12. Design Council (n.d.) History of the Double Diamond. https://www.designcouncil.org.uk/our-resources/the-double-diamond/history-of-the-double-diamond/
13. EIT. Creation of a Taxonomy for the European AI Ecosystem (2021). https://eit.europa.eu/library/creation-taxonomy-european-ai-ecosystem
14. Eshet-Alkali, Y.: Digital literacy: a conceptual framework for survival skills in the digital era. J. Educ. Multimedia Hypermedia **13**(1), 93–106 (2004)
15. European Commission. AI Act (2024). https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
16. Fischer, L., Ehrlinger, L., Geist, V., et al.: AI system engineering—key challenges and lessons learned. Mach. Learn. Knowl. Extr. **3**, 56–83 (2021). https://doi.org/10.3390/make3010004
17. Gilster, P.: Digital Literacy. Wiley Computer Publishing (1997)
18. Goetz, L., Seedat, N., Vandersluis, R., et al.: Generalization—A key challenge for responsible AI in patient-facing clinical applications. NPJ Digit. Med. **7**(126) (2024). https://doi.org/10.1038/s41746-024-01127-3
19. Hartikainen, M., Väänänen, K., Lehtiö, A., et al.: Human-Centered AI design in reality: a study of developer companies' practices. nordic human-computer interaction conference (NordiCHI '22). ACM **55**, 1–11 (2022). https://doi.org/10.1145/3546155.3546677
20. Hartikainen, M., Väänänen, K., and Olsson, T.: Towards a human-centred artificial intelligence maturity model. In: CHI Conference on Human Factors in Computing Systems (CHI EA '23). ACM, vol. 285, pp. 1–7 (2023). https://doi.org/10.1145/3544549.3585752
21. Julien, H.: Digital literacy in theory and practice. In: Khosrow-Pour, M. (ed.). Encyclopedia of Information Science and Technology, Fourth Edition, pp. 2243–2252. IGI Global Scientific Publishing (2018). https://doi.org/10.4018/978-1-5225-2255-3.ch195
22. Kiernan, L., Ledwith, A., Lynch, R.: An exploration of the cognitive processes of design teams to inform design education and practice. Proc. Design Soc. **2**, 1235–1244 (2022). https://doi.org/10.1017/pds.2022.126
23. Kneusel, R.T.: How AI works: from sorcery to science. No Starch Press (2024)
24. Lindenfalk, B., Resmini, A.: The dialectic between system space and design space. In: Soares, M.M., Rosenzweig, E., Marcus, A. (eds.) Design, User Experience, and Usability: Design

Thinking and Practice in Contemporary and Emerging Technologies. HCII 2022. LNCS, 13323. Springer (2022). https://doi.org/10.1007/978-3-031-05906-3_3
25. Long, D., Magerko, B.: What is AI literacy? Competencies and design considerations. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–16 (2020). https://doi.org/10.1145/3313831.3376727
26. Lu, Q., Zhu, L., Xu, X., et al.: A taxonomy of foundation model based systems through the lens of software architecture. In: IEEE/ACM 3rd International Conference on AI Engineering – Software Engineering for AI (CAIN), pp. 1–6 (2024)
27. Lutnæs, E.: Framing the concept design literacy for a general public. Form Akademisk, Res. J. Design Design Educ. **14**(4) (2021). https://doi.org/10.7577/formakademisk.4639
28. Narayanan, A., Kapoor, S.: AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference. Princeton University Press, Princeton (2024)
29. Ng, W.: Can we teach digital natives' digital literacy? Comput. Educ. **59**(3), 1065–1078 (2012). https://doi.org/10.1016/j.compedu.2012.04.016
30. Ng, W., Leung, M.Y., Chan, W.: AI literacy in the digital age: understanding, evaluating, and interacting with AI technologies. Comput. Educ. **172**, 104252 (2021). https://doi.org/10.1016/j.compedu.2021.104252
31. Pacione, C.: Evolution of the mind: a case for design literacy. Interactions **17**(2), 6–11 (2010). https://doi.org/10.1145/1699775.1699777
32. Pilgrim, J., Martinez, E.: Defining literacy in the 21st century: a guide to terminology and skills. Texas J. Literacy Educ. **1**(1), 60–69 (2013). https://doi.org/10.1016/j.compedu.2021.104252
33. Quintarelli, E., Resmini, A., Rosati, L.: FaceTag: integrating bottom-up and top-down classification in a social tagging system. J. Bull. Am. Soc. Inf. Sci. Technol. **33**(5), 10–15 (2007)
34. Ranganathan, S.R.: Colon classification. Asia Publishing House (1933)
35. Resmini, A.: Information architecture in the age of complexity. Bull. Am. Soc. Inf. Sci. Technol. (2012). https://doi.org/10.1002/bult.2012.1720390104
36. Resmini, A., Rosati, L.: Pervasive Information Architecture. Elsevier, pp. 157–158 (2011)
37. Resmini, A., Rosati, L.: The semantic environment: heuristics for a cross-context human–information interaction model. In: Dubois, E., Gray, P., Nigay, L. (eds.) The Engineering of Mixed Reality Systems. HCI Series. Springer (2010). https://doi.org/10.1007/978-1-84882-733-2_5
38. Riedl, M.: Human-centered artificial intelligence and machine learning. Hum. Behav. Emerging Technol. Special Issue Article (2019). https://doi.org/10.1002/hbe2.117
39. Rousse, B.S. Dreyfus, S.E.: Revisiting the six stages of skill acquisition. In Silva Mangiante, E., Peno, K. (eds.) Teaching and Learning for Adult Skill Acquisition: Applying the Dreyfus & Dreyfus Model in Different Fields. Information Age Publishing (2021)
40. Shneiderman, B.: Human-Centered AI. Oxford University Press, Oxford (2022)
41. Siegel, E.: Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die. Wiley, Hoboken (2013)
42. Smith, G.: The AI Delusion. Oxford University Press, Oxford (2018)
43. Windl, M., Feger, S.S., Zijlstra, L. et al.: It is not always discovery time: Four pragmatic approaches in designing AI systems. In: CHI Conference on Human Factors in Computing Systems (CHI '22). ACM (2022). https://doi.org/10.1145/3491102.3501943
44. Yang, Q., Steinfeld, A., Rosé, C., Zimmerman, J.: Re- examining whether, why, and how human-AI interaction is uniquely difficult to design. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, pp. 1–13. ACM (2020). https://doi.org/10.1145/3313831.3376301

Generative AI in Physics Education: Evaluating Tools and Teaching Approaches

Keyi Qiu[1(✉)] and Tianrun Zhang[2]

[1] Dublin High School, Dublin, CA, USA
keyiqiu1397@gmail.com
[2] Zhejiang University, Hangzhou, China
3210101824@zju.edu.cn

Abstract. This study explores how Generative AI (GenAI) tools—ChatGPT, Bing Copilot, and Google Gemini—impact learning and engagement in university-level physics education. Merging systematic literature review and classroom experimentation, we evaluate how GenAI supports conceptual understanding, problem-solving, and user experience. Our findings indicate that GenAI enables flexible, personalized learning by simplifying complex topics and providing real-time feedback. Distinct affordances emerged across platforms: ChatGPT for logical reasoning, Bing Copilot for procedural tasks, and Google Gemini for visual explanations. However, limitations persist, including AI's inability to address deep conceptual misconceptions and the risk of over-reliance. From an HCI perspective, this work underscores the importance of designing AI-powered educational tools that encourage active learning while aligning with human cognitive needs. We argue that careful integration and instructor mediation are critical for maximizing learning outcomes. This research contributes to the design of human-centered AI systems that enhance STEM education through adaptive, responsive, and interactive learning environments.

Keywords: Generative AI · Human-Computer Interaction · Physics Education · Adaptive Learning · ChatGPT · Bing Copilot · Google Gemini · Conceptual Understanding · AI-assisted learning · Educational Technology

1 Introduction

Generative AI (GenAI) is transforming education, especially in STEM fields like physics. Tools such as ChatGPT, Bing Copilot, and Google Gemini offer personalized, real-time support, enhancing students' grasp of abstract concepts and problem-solving skills in university-level physics.

These technologies simplify complex content, generate practice problems, and provide step-by-step solutions, reshaping traditional physics education. GenAI

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 397–405, 2025.
https://doi.org/10.1007/978-3-031-94171-9_38

empowers students to explore principles independently and promotes flexible learning, allowing them to progress at their own pace through repetition and feedback.

While GenAI shows promise, its integration and effectiveness in physics learning require further investigation. This study examines GenAI's practical applications, impact on learning outcomes, and platform limitations.

1.1 Research Questions

This research addresses the following questions:

- How can GenAI tools assist university students in understanding physics concepts?
- What are the most effective strategies for integrating GenAI into the physics curriculum?
- Which GenAI platforms offer the greatest benefits, and what are their limitations?

2 Related Works

Generative AI is influencing education by providing tools for personalized learning and assessment. Several studies have examined how large language models (LLMs), including ChatGPT and Gemini, explain physics concepts like the conservation of momentum, showcasing their strengths and weaknesses in different settings [1,3,5,6,9–16,18,20]. ChatGPT offers more nuanced explanations, while Gemini is more accessible for beginners. These studies highlight the need for educator guidance to help students fully utilize these tools [2].

Several studies have examined how AI tools influence conceptual understanding in physics education. Prior work, such as that of Domenichini et al. [4], explores the integration of gamified virtual environments and personalized AI-generated learning activities to improve engagement and learning outcomes. Inspired by this, our study evaluates whether Generative AI tools can similarly enhance student engagement and comprehension in university-level physics courses.

A review of GenAI in education identified its ability to provide personalized learning experiences and assist in content generation across disciplines. In physics, students have used ChatGPT to gain deeper insights into topics like classical mechanics and electromagnetism [8]. However, ethical concerns, such as biases in AI-generated content, underscore the need for careful integration into classrooms [19].

A study on ChatGPT in physics demonstrated how simple activities fostered critical thinking and engagement [17]. In this pilot, ChatGPT generated practice problems and guided students through problem-solving. Educators, however, stressed the need to balance AI use with independent problem-solving skills.

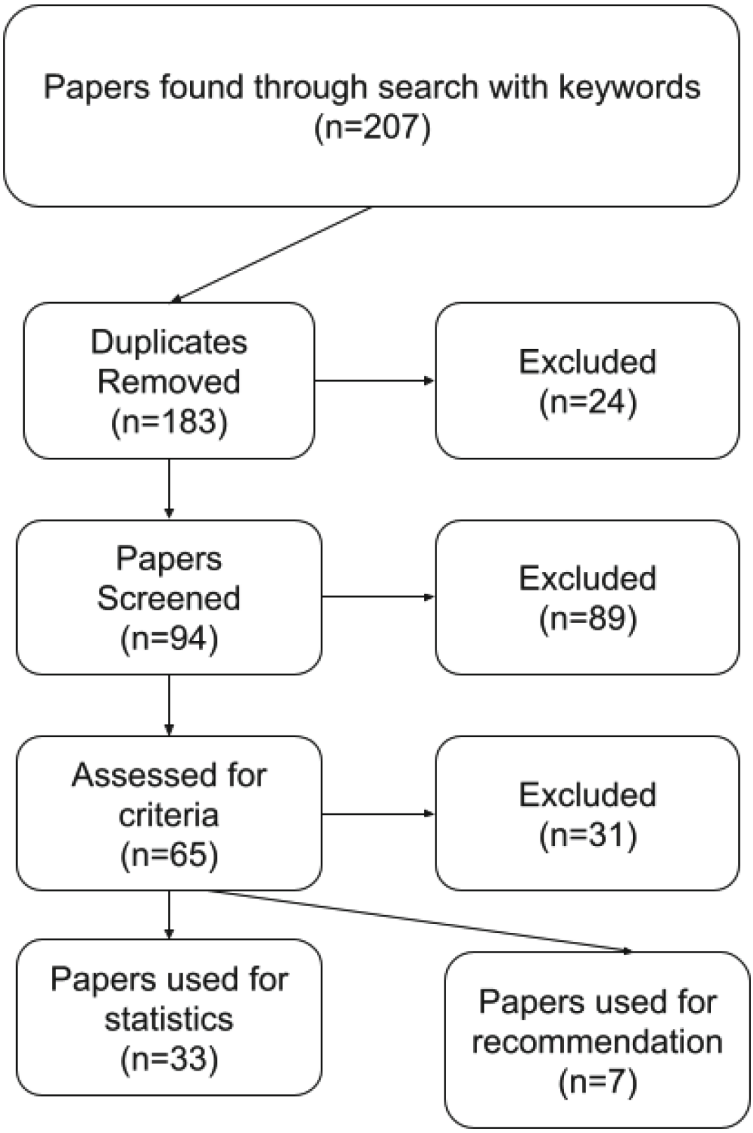


Fig. 1. Selection process for the papers.

AI-assisted grading systems have been explored to address the time-consuming task of grading complex problems. GenAI offers formative feedback and reduces grading time, particularly in large classes [21], but concerns remain about the reliability of AI feedback, especially with nuanced student responses.

Despite these advancements, integrating GenAI into physics education poses challenges. Over-reliance on AI may disengage students from critical thinking,

and AI can generate misleading solutions. A balanced approach between AI-assisted and traditional methods is essential. Further research is needed to fully assess GenAI's potential.

3 Methods

This study follows the PRISMA approach to systematically review existing research on the role of Generative AI (GenAI) in physics education while also conducting ongoing experiments to evaluate its effectiveness in university-level physics learning. The literature review focused on research published between 2019 and 2024, primarily selecting studies that examined the application of AI tools in college-level physics education. In cases where college-level data was limited, relevant studies on high school students were included. The search, conducted through Google Scholar and arXiv.org, initially identified 207 papers, which were filtered down to 65 studies after removing duplicates and applying exclusion criteria. Studies without empirical data or those that did not directly address AI's impact on physics education were excluded.

3.1 Study Selection

To assess the effectiveness of AI platforms, this study focused on three key aspects: conceptual understanding, student engagement, and problem-solving assistance. Previous research often measured conceptual understanding by analyzing students' ability to connect different physics principles rather than simply recalling formulas. Several studies employed the Taxonomy of Introductory Physics Problems (TIPP) to categorize AI-generated explanations based on their complexity, a method we adopted in our evaluation. In particular, we examined whether AI tools like ChatGPT, Bing Copilot, and Google Gemini facilitated deeper learning by guiding students through synthesis problems and context-rich scenarios that required them to integrate multiple physics concepts.

The literature search, conducted via Google Scholar and arXiv.org, initially identified 207 studies. Keywords included "Generative AI in physics education", "ChatGPT for physics learning", and "AI-supported learning in STEM". After removing duplicates, 183 papers remained. Based on our selection criteria, we excluded an additional 89 papers. For the remaining 94 papers, we carefully assessed their relevance and quality, narrowing the list to 65. Among these, 33 papers with clear data were included for statistical analysis. Finally, we recommend seven papers as potential best practices, as shown in Fig. 1.

3.2 Data Collection and Evaluation

Key data, such as performance evaluations, benefits, and limitations of GenAI in physics education, were extracted from the selected studies. GenAI platforms were assessed for their role in physics learning tasks like generating explanations, solving problems, and providing feedback. This review synthesized initial

findings, while further meta-analysis will be conducted in future work to evaluate long-term efficacy. The thematic analysis compared how each study applied GenAI, focusing on practical applications and challenges.

Prior research has explored the potential of generative AI models in physics education, particularly in evaluating students' conceptual understanding through standardized assessments. Kieser et al. [7] examined ChatGPT's ability to solve the Force Concept Inventory (FCI), a widely used concept test for Newtonian mechanics. Their findings suggest that AI models can accurately solve structured physics problems but may struggle with nuanced conceptual reasoning. Inspired by this, our study evaluates how ChatGPT, Bing Copilot, and Google Gemini support problem-solving and conceptual understanding in an introductory college-level physics course.

Additionally, recent studies have investigated AI-assisted grading for handwritten student solutions. Kortemeyer [8] explored a feasibility study on using AI to evaluate handwritten physics problem derivations. The study found that tools like MathPix and GPT-4 could scan and grade student solutions with a high correlation to human graders (R-squared of 0.84). While AI-assisted grading appears promising for formative feedback, its reliability for summative assessments remains uncertain. This aligns with our interest in assessing AI's role not only in conceptual learning but also in evaluating students' problem-solving strategies and accuracy.

Beyond conceptual understanding, the study also considered AI's impact on student engagement and motivation. Prior work suggested that gamified AI-driven learning environments could significantly enhance motivation, particularly for students who struggle with traditional instruction. While our research did not directly implement a gamification framework, we examined student perceptions of AI's role in maintaining engagement. Student feedback was collected to understand whether AI-assisted explanations helped sustain interest and improve confidence in problem-solving. Instructor evaluations also played a role in assessing how effectively AI-generated explanations complemented traditional teaching methods.

A significant part of this study also involved analyzing AI's problem-solving assistance. Given that physics education relies heavily on problem-based learning, we compared AI-generated solutions with those provided by human instructors. This comparison focused on whether AI-generated solutions were logically structured, easy to follow, and free of errors. Since Newtonian mechanics is often considered a foundational topic in physics, it was chosen as the primary area of focus when assessing AI performance. Additionally, this study examined whether AI-assisted learning paths could adapt to individual student needs by tailoring explanations and feedback based on the complexity of the problems attempted.

3.3 Ongoing Experiments

In addition to the literature review, ongoing experiments are being conducted in an introductory college-level physics course. These experiments evaluate how GenAI supports problem-solving and conceptual understanding. Preliminary

results indicate that AI-powered platforms such as ChatGPT and Bing Copilot help students break down complex physics principles into more digestible components. Google Gemini, in particular, has been effective for students who benefit from visual learning, as it provides intuitive diagrams and graphical explanations alongside textual descriptions. The study further examines how AI-generated explanations support fundamental problem-solving strategies, particularly in Newtonian mechanics.

As part of the experiment, we are also investigating AI's potential role in automated feedback and grading. Inspired by prior work on AI-assisted assessment [8], we explore whether AI-generated feedback aligns with human grading criteria and whether it can assist students in recognizing conceptual misunderstandings. This aspect of the study will provide insights into the feasibility of using AI-driven evaluation tools in physics education.

Since this research involved human participants, Institutional Review Board (IRB) approval was obtained before data collection. All participants provided informed consent, and their responses were anonymized to ensure privacy. While a full explanation of ethical compliance measures could not be included due to space constraints, additional details are available upon request.

4 Results

Preliminary findings suggest that Generative AI tools provide significant benefits in enhancing conceptual understanding and problem-solving skills in physics education. These findings address the research questions as follows:

1. GenAI tools simplify complex concepts and provide personalized feedback, helping students develop a deeper understanding of physics principles.
2. Effective use of these tools involves integrating them as supplementary aids rather than primary instructional methods to maintain active student engagement.
3. ChatGPT excels in logical reasoning and step-by-step explanations, Bing Copilot is more effective for procedural calculations, and Google Gemini offers strong accessibility features. However, AI over-reliance remains a concern.

The study also examined how AI influences student engagement and motivation. Many students reported that AI-assisted learning helped them feel more confident in approaching complex problems, especially when AI provided step-by-step feedback. Similar to findings from prior studies on AI-driven gamification, some students expressed that AI responses made physics feel less intimidating and more interactive. However, one emerging concern was over-reliance on AI-generated answers. Some students admitted to passively accepting AI solutions without fully engaging in independent problem-solving. This highlights an important pedagogical challenge: while AI can support learning, it must be integrated in a way that encourages active critical thinking rather than passive consumption of information.

Our findings also align with prior research suggesting that AI-generated explanations sometimes fail to fully capture conceptual misunderstandings in physics, particularly in assessments like the Force Concept Inventory (FCI) [7]. While AI models provide strong procedural support, they occasionally lack the ability to correct students' deeper misconceptions, reinforcing the need for human oversight in AI-assisted learning.

Beyond conceptual explanations, AI tools also played a role in adaptive learning. In comparing ChatGPT and Bing Copilot, it was observed that ChatGPT was more effective for logical reasoning and step-by-step explanations, whereas Bing Copilot performed better for solving procedural calculations. However, AI-generated feedback was less reliable when dealing with open-ended reasoning tasks, particularly when students were required to justify their answers conceptually rather than simply providing numerical results. Instructor evaluations indicated that human oversight remains essential when using AI tools for learning, as AI occasionally provided responses that were overly generalized or lacked the depth required for higher-level problem-solving.

Furthermore, our analysis of AI-generated feedback aligns with the findings of Wan and Chen [19], who reported that AI-generated feedback was well-received by students, particularly in cases requiring conceptual clarification. In our study, students generally found AI-generated feedback helpful in breaking down problem-solving steps, but limitations emerged when addressing highly conceptual or ambiguous problems. This reinforces the idea that AI can be an effective supplemental tool but may not fully replace instructor-led feedback in physics education.

5 Discussion

While our findings highlight the potential of Generative AI tools in physics education, several limitations should be acknowledged. AI-generated explanations, though useful for structured problem-solving, sometimes fail to address deeper conceptual misunderstandings, aligning with prior research [7]. Additionally, student over-reliance on AI remains a concern, as some learners rely on AI-generated solutions rather than engaging in independent reasoning.

Recent studies have also explored the effectiveness of AI-generated feedback in supporting conceptual learning. Wan and Chen [19] found that while students appreciated AI-generated feedback for its clarity and accessibility, there were limitations in its ability to address deeper misconceptions. Our findings reinforce this, as AI-generated feedback in our study was most effective when clarifying procedural errors but sometimes fell short in explaining broader conceptual issues.

Future research should explore how AI can be better integrated into instructional design to balance automation with active learning. More robust evaluation frameworks are needed to assess long-term impacts on student performance. Expanding experiments to include a broader range of physics topics and incorporating instructor-led AI interventions may further improve the effectiveness of these tools.

6 Conclusion

This study examined the role of Generative AI platforms in enhancing conceptual understanding and problem-solving in physics education. Through a systematic review and ongoing experiments, we found that ChatGPT, Bing Copilot, and Google Gemini assist in simplifying complex concepts, improving engagement, and providing adaptive support. However, challenges such as AI dependency and conceptual depth limitations highlight the need for careful integration. Future studies should refine AI-based instructional strategies to maximize learning benefits while mitigating potential drawbacks.

References

1. Allen-Zhu, Z., Li, Y.: Physics of language models: Part 3. 1, knowledge storage and extraction. arXiv preprint arXiv:2309.14316 (2023)
2. Bitzenbauer, P.: ChatGPT in physics education: a pilot study on easy-to-implement activities. Contemp. Educ. Technol. **15**(3), ep430 (2023)
3. Ding, L., Li, T., Jiang, S., Gapud, A.: Students' perceptions of using ChatGPT in a physics class as a virtual tutor. Int. J. Educ. Technol. High. Educ. **20**(1), 63 (2023)
4. Domenichini, D., Bucchiarone, A., Chiarello, F., Schiavo, G., Fantoni, G.: An AI-driven approach for enhancing engagement and conceptual understanding in physics education. In: 2024 IEEE Global Engineering Education Conference (EDUCON), pp. 1–3. IEEE (2024)
5. Holmes, J., et al.: Evaluating large language models on a highly-specialized topic, radiation oncology physics. Front. Oncol. **13**, 1219326 (2023)
6. Kasneci, E., et al.: ChatGPT for good? On opportunities and challenges of large language models for education. Learn. Individ. Differ. **103**, 102274 (2023)
7. Kieser, F., Wulff, P., Kuhn, J., Küchemann, S.: Educational data augmentation in physics education research using ChatGPT. Phys. Rev. Phys. Educ. Res. **19**(2), 020150 (2023)
8. Kortemeyer, G.: Toward AI grading of student problem solutions in introductory physics: a feasibility study. Phys. Rev. Phys. Educ. Res. **19**(2), 020163 (2023)
9. Li, W., Zhou, N., Qu, X.: Enhancing eye-tracking performance through multi-task learning transformer. In: International Conference on Human-Computer Interaction, pp. 31–46. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-61572-6_3
10. Liang, Y., Zou, D., Xie, H., Wang, F.L.: Exploring the potential of using ChatGPT in physics education. Smart Learn. Environ. **10**(1), 52 (2023)
11. Mittal, U., Sai, S., Chamola, V., et al.: A comprehensive review on generative AI for education. IEEE Access (2024)
12. Murungi, N.K., Pham, M.V., Dai, X.C., Qu, X.: Empowering computer science students in electroencephalography (EEG) analysis: a review of machine learning algorithms for EEG datasets. In: The 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) (2023)
13. Polverini, G., Gregorcic, B.: How understanding large language models can inform the use of ChatGPT in physics education. Eur. J. Phys. **45**(2), 025701 (2024)

14. Qu, X., Key, M., Luo, E., Qiu, C.: Integrating HCI datasets in project-based machine learning courses: a college-level review and case study. In: Degen, H., Ntoa, S. (eds) International Conference on Human-Computer Interaction, pp. 124–143. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-76827-9_8
15. Sato, K.: Exploring the educational landscape of AI: large language models' approaches to explaining conservation of momentum in physics. arXiv preprint arXiv:2407.05308 (2024)
16. Saunders, T., Aleisa, N., Wield, J., Sherwood, J., Qu, X.: Optimizing the literature review process: evaluating generative AI models on sum-marizing undergraduate data science research papers. In: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (2024)
17. Suhonen, S.: Navigating the AI era: analyzing the impact of artificial intelligence on learning and teaching engineering physics. In: Proceedings of the 12th International Conference on Physics Teaching in Engineering Education PTEE 2024, pp. 76–83 (2024)
18. Tong, D., et al.: Investigating ChatGPT-4's performance in solving physics problems and its potential implications for education. Asia Pacific Educ. Rev., 1–11 (2023)
19. Wan, T., Chen, Z.: Exploring generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning. Phys. Rev. Phys. Educ. Res. **20**(1), 010152 (2024)
20. Yunoki, I., Berreby, G., D'Andrea, N., Lu, Y., Qu, X.: Exploring AI music generation: a review of deep learning algorithms and datasets for undergraduate researchers. In: Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (eds) International Conference on Human-Computer Interaction, pp. 102–116. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-49215-0_13
21. Yusuf, A., Pervin, N., Román-González, M., Noor, N.M.: Generative AI in education and research: a systematic mapping review. Rev. Educ. **12**(2), e3489 (2024)

Cognitive Flexibility and Attitude Toward AI: A Correlational Study

Andrés Subía-Arellano[1], Doris Pérez-Vega[1], Rocio Patiño-Fernández[2], and Jorge Buele[3](✉)

[1] Facultad de Ciencias Sociales y Humanas, Universidad Tecnológica Indoamérica, Quito 170103, Ecuador
[2] Departamento de Idiomas, Universidad Tecnológica Indoamérica, Quito 170103, Ecuador
[3] Centro de Investigación en Mecatrónica y Sistemas Interactivos (MIST), Facultad de Ingenierías, Universidad Tecnológica Indoamérica, Ambato 180103, Ecuador
jorgebuele@uti.edu.ec

Abstract. Artificial intelligence plays a leading role across various sectors, underscoring the importance of understanding the individual factors influencing their acceptance. Previous research has pointed out that variables such as age, gender, and cognitive flexibility impact attitudes toward these technologies. However, the interaction among these variables still requires further analysis. This study sought to explore the relationships between cognitive flexibility, age, gender, and attitudes toward artificial intelligence in a sample of 342 participants, with an average age of 26.80 years. Employing a descriptive-correlational design, two scales were used: one to measure cognitive flexibility and another to assess attitudes toward this technology. Due to the lack of normality in the distributions of the variables, Spearman's correlation was used for the analysis. The results show that cognitive flexibility and educational level have a positive and significant relationship with the attitude toward artificial intelligences ($r = 0.245$, $p < 0.001$ and $r = 0.140$, $p = 0.009$, respectively). On the other hand, age presents a weak negative relationship ($r = -0.117$, $p < 0.05$), while no significant relationship was observed with gender. These findings provide an initial basis for understanding individual differences in technology acceptance, although further research is needed to delve into the underlying mechanisms and evaluate other contextual factors.

Keywords: Cognitive Flexibility · Artificial Intelligence · Technological Acceptance

1 Introduction

Over the years, various technologies such as big data, virtual reality, and the internet of things (IoT) have transformed industries and daily life, optimizing processes and enhancing efficiency [1–3]. More recently, artificial intelligence (AI) has gained a leading role in various areas of society, from automation in industry to decision-making in key sectors such as health and education [4, 5]. However, individuals' attitudes toward these technologies vary depending on factors such as age, gender, and cognitive abilities

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 406–411, 2025.
https://doi.org/10.1007/978-3-031-94171-9_39

[6]. Understanding these attitudes is important, as it can influence the successful implementation of AI-based technologies in different contexts [7]. Among cognitive abilities, cognitive flexibility has been identified as a key variable in adapting to technological changes and innovations involving AI [8].

This skill, defined as the ability to adapt to new situations and switch perspectives, enables individuals to adjust their thoughts and behaviors in the face of constantly evolving technological environments [9]. Those with higher cognitive flexibility may have a more receptive attitude toward AI, facilitating their acceptance and use in everyday life [8]. Moreover, demographic factors such as age have been shown to be determinants in technology acceptance. Previous studies have indicated that older individuals tend to have a lower readiness to adopt new technologies due to limitations in familiarity and technological confidence [10]. On the other hand, recent research has explored how gender and educational level also influence attitudes toward AI, though with mixed results [reference]. The present study aims to explore the relationships between cognitive flexibility, age, gender, and attitude toward artificial intelligence. Through a correlational analysis, it seeks not only to identify how these factors interact but also to understand the barriers and facilitators in the acceptance of AI-based technologies [11]. This approach will provide evidence to design strategies that promote more inclusive adoption of AI across various population groups.

2 Methodology

2.1 Study Design

This study employs a descriptive-correlational design to analyze the relationship between cognitive flexibility and attitude toward artificial intelligence in a sample of university participants.

2.2 Population

The study sample consisted of 342 participants, ranging in age from 17 to 67 years (M = 26.80, SD = 9.552). Of these, 133 (38.9%) were male and 209 (61.1%) female. The gender distribution shows a higher female representation. These sociodemographic data provide a general overview of the diversity of participants, allowing for the analysis of the relationship between variables across different age and gender groups. The sample selection was non-probabilistic and based on the availability of participants.

2.3 Instruments

Two validated scales were used for data collection: Cognitive Flexibility Scale (CFS) by Martin and Rubin (1995). This scale assesses the ability to perceive multiple alternatives and adapt to changes, with 12 items measuring three dimensions: Awareness of multiple options, Willingness to be flexible, and Confidence in the ability to adapt. Its reliability, measured through Cronbach's Alpha, was 0.733, indicating good internal consistency. EVAIA-1 Scale (Design of the Attitudinal Assessment Scales towards Artificial Intelligence) (Subía, Pérez, Guillén & Cáceres, 2023), consisting of 12 items designed to

measure attitudes toward artificial intelligence in various contexts. The scale showed high reliability (Cronbach's Alpha = 0.798).

2.4 Procedure

The instruments were administered both in-person and virtually, ensuring the confidentiality and anonymity of the responses. The purpose of the study was explained, and informed consent was obtained from each participant before administering the questionnaires. Instructions were ensured to be clear and uniform.

2.5 Data Analysis

For data analysis, descriptive and correlational statistics were used. Measures of central tendency (mean, median) and dispersion (standard deviation) were analyzed. Subsequently, to assess the relationship between variables of cognitive flexibility, working memory, and attitude toward artificial intelligence, Spearman's correlation test was applied, due to the ordinal nature of the scales used.

3 Results

3.1 Scale Reliability

The internal consistency of the cognitive flexibility scale was evaluated using Cronbach's Alpha coefficient. The results indicate good reliability of the scale, with a Cronbach's Alpha of 0.733, which exceeds the minimum acceptable threshold of 0.70. Additionally, the Alpha based on standardized items was 0.762. The scale consists of 12 items, which coherently measure the construct of cognitive flexibility. The reliability of the attitude toward artificial intelligences scale was evaluated using the Cronbach's Alpha coefficient, as shown in Table 2. The result obtained was 0.798, indicating good internal consistency. This value exceeds the acceptable threshold of 0.70, confirming that the 12 items comprising the scale coherently measure the construct of attitude toward artificial intelligences.

3.2 Correlations

The normality of the distributions was evaluated using the Kolmogorov-Smirnov test. The results indicated that both the Total Cognitive Flexibility variable ($D = 0.068$, $p = 0.001$) and the Total Attitude toward AI variable ($D = 0.070$, $p < 0.001$) exhibit significant deviations from normality. Since both significance values (p) are less than 0.05, it is concluded that none of the variables follow a normal distribution. Consequently, to examine the relationships between sociodemographic, cognitive variables, and attitude toward AI, Spearman's correlation test was used, which is appropriate for data that do not meet the assumption of normality. Below are the results of the correlations and their interpretation.

The correlation table (Table 1) shows interesting results about the relationships between attitude toward artificial intelligences (AI) and sociodemographic and cognitive variables. A weak but significant negative correlation was observed with age ($r = -0.117$, $p < 0.05$). This suggests that older individuals have a less favorable attitude toward AI, which could be explained by less exposure and familiarity with these technologies compared to younger generations, who have grown up in a more advanced technological environment.

Table 1. Correlations between sociodemographic variables, cognitive flexibility, and attitude toward AI.

Variable	Age	Gender	University studies	Total Cognitive Flexibility
Total attitude toward AI	-0.117*	< 0.001	0.140**	0.245**
Significance	0.030	0.998	0.009	< 0.001

* The correlation is significant at the 0.05 level.
** The correlation is significant at the 0.01 level.

On the other hand, no significant relationship was found between attitude toward AI and gender ($r = 0.000$, $p = 0.998$). This result indicates that, regardless of gender, attitudes toward AI are similar, reflecting a neutral and homogeneous perception of the impact of this technology in different contexts. In contrast, a significant positive correlation was identified with the level of university studies ($r = 0.140$, $p < 0.01$). This finding may be due to the fact that individuals with higher educational levels tend to have more knowledge and understanding of the benefits and applications of AI, fostering more positive attitudes. Finally, the strongest correlation was observed with cognitive flexibility ($r = 0.245$, $p < 0.01$), highlighting the importance of adaptability and openness to change in the acceptance of emerging technologies. Individuals with higher cognitive flexibility tend to be more receptive to new ideas and less resistant to innovations, which explains their favorable attitude toward AI. Together, these results underscore the relevance of factors such as age, educational level, and cognitive flexibility in shaping attitudes toward advanced technologies. This has implications for the implementation of technology literacy programs, especially targeted at older populations or those with lower educational levels, to encourage greater acceptance and utilization of AI in various fields.

4 Conclusions

The present study reveals that cognitive flexibility and educational level are key factors in fostering a positive attitude toward AI, while age acts as a barrier in its acceptance. These findings reinforce the idea that individual differences influence how emerging technologies are perceived and adopted. Additionally, the absence of a significant relationship between gender and attitude toward AI suggests that technological perception may be independent of gender, opening the possibility for generalized approaches in certain educational strategies.

However, the study has some limitations. The sample was selected for convenience, limiting the generalization of the results to other populations. Moreover, the cross-sectional design does not allow for establishing causal relationships between the analyzed variables. For future research, it is recommended to expand the sample to include a greater demographic diversity and use longitudinal designs to explore how attitudes toward AI evolve over time and in different usage contexts. It would also be valuable to incorporate additional variables, such as previous experience with AI and the level of technological literacy, to obtain a more comprehensive picture of the factors influencing the acceptance of these technologies.

Acknowledgments. This study was supported by Universidad Tecnológica Indoamérica under project IIDI-007–25 titled "La respuesta psicofisiológica frente a estimulación multisensorial en personas empáticas y agresivas".

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Campaña-Cordova, M., Castillo-Ledesma, F., Oleas-Orozco, J., Moncayo, H., Avilés-Castillo, F.: Enhancing the teaching of natural sciences in rural environments through educational videos. In: 2023 IEEE Seventh Ecuador Technical Chapters Meeting (ECTM), pp. 1–5 (2023). https://doi.org/10.1109/ETCM58927.2023.10309022
2. Buele, J., Avilés-Castillo, F., Del-Valle-Soto, C., Varela-Aldás, J., Palacios-Navarro, G.: Effects of a dual intervention (motor and virtual reality-based cognitive) on cognition in patients with mild cognitive impairment: a single-blind, randomized controlled tri-al». J. NeuroEng. Rehabil. **21**(1), 130 (2024). https://doi.org/10.1186/s12984-024-01422-w
3. Avilés-Castillo, F., Ayala-Chauvin, M., Buele, J.: Evolution, collaborations, and impacts of big data research in Ecuador: bibliometric analysis. In: Guarda, T., Portela, F., Diaz-Nafria, J.M. (eds.) Advanced Research in Technologies, Information, In-novation and Sustainability, pp. 290–301. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-48855-9_22
4. Mogi, K.: Artificial intelligence, human cognition, and conscious supremacy. Front. Psychol. **15** (2024). https://doi.org/10.3389/fpsyg.2024.1364714
5. Matthews, G., et al.: Evolution and revolution: Personality research for the coming world of robots, artificial intelligence, and autonomous systems. Personal. In-divid. Differ. **169**, 109969 (2021). https://doi.org/10.1016/j.paid.2020.109969
6. Subía-Arellano, A., Pérez-Vega, D., Guillen-García, S., Cáceres-Fierro, N.: Design of the attitudinal assessment scale towards artificial intelligence (EVAIA-1). In: 2023 IEEE Seventh Ecuador Technical Chapters Meeting (ECTM), pp. 1–6 (2023). https://doi.org/10.1109/ETCM58927.2023.10309057
7. Sabando-García, Á.R., Moreira-Choez, J.S., Bazurto-Loor, A.A., Bravo-Madrid, A.B.: Alfabetización en inteligencia artificial en modalidades de estudio presencial y virtual entre estudiantes universitarios. Soc. Educ. Lens. **1** (2024). https://doi.org/10.56931/sel.2024.e4
8. Kelly, S., Kaye, S.A., Oviedo-Trespalacios, O.: What factors contribute to the ac-ceptance of artificial intelligence? A systematic review. Telemat. Inf. **77** (2023). https://doi.org/10.1016/j.tele.2022.101925

9. Lent, R.W., Brown, S.D.: Social cognitive model of career self-management: toward a unifying view of adaptive career behavior across the life span. J. Couns. Psychol. **60**, 557–568 (2013). https://doi.org/10.1037/a0033446
10. Samuel, J., Kashyap, R., Samuel, Y., Pelaez, A.: Adaptive cognitive fit: artificial intelligence augmented management of information facets and representations. Int. J. Inf. Manag. **65**, 102505 (2022). https://doi.org/10.1016/j.ijinfomgt.2022.102505
11. Mikalef, P., Gupta, M.: Artificial intelligence capability: conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Inf. Manage. **58**, 103434 (2021). https://doi.org/10.1016/j.im.2021.103434

The Impact of Dialogue with ChatGPT and Extroversion on Mental State

Shunichi Tahara[1(✉)], Eriko Sugisaki[1], Keiji Yasuda[1], Akihiro Miyamoto[2], Motoyuki Sanada[2], Kae Nakajima[2], Kazushi Ikeda[1], Nao Kobayashi[1], and Yasushi Naruse[2]

[1] KDDI Research, Inc., 2-1-15, Ohara, Fujimino-Shi, Saitama, Japan
sh-tahara@kddi.com

[2] National Institute of Information and Communications Technology, 588-2, Iwaoka, Iwaoka-Cho, Nishi-Ku, Kobe-Shi, Hyogo, Japan

Abstract. Conversations with others are important in order to improve an individual's mental state. Recently, LLM-based dialogue systems, such as ChatGPT, facilitate natural conversation and have the potential to improve mental state such as emotions or mood. Our aim is to develop a dialogue system that generates responses tailored to the user's personality and physiological state. As part of this study, we set the research question *"Does dialogue with ChatGPT, and the interplay of the extroversion between the user and ChatGPT, improve the user's mental state?"*. In order to verify this research question, we implemented a dialogue application, and conducted a total of three speech dialogue experiments using a ChatGPT API. Participants rated their mental state along seven emotion axes (e.g., Sad-Happy, Anxious-Relaxed), using a 15-point scale both before and after each dialogue. To set the extroversion level (low, average or high) for ChatGPT, we added information of ChatGPT's extroversion into a prompt. Our experiments showed that the mean of seven emotion axes significantly improved after dialogues. Moreover, although not statistically significant, the largest differences between before and after dialogue were observed when both participants and ChatGPT had either a similarly low or high level of extroversion.

Keywords: ChatGPT · Mental State · Dialogue System

1 Introduction

In daily life, conversations with others are important in order to improve an individual's mental state. For example, Shakya et al. found that a person's mental health can be improved through conversations between friends [1]. However, there are specific situations where conversations with others may be difficult. For instance, athletes who wish to boost their mood right before a match might not have anyone around to converse with, resulting in the inability to improve their mood.

Recently, LLM-based dialogue systems such as ChatGPT, could facilitate natural conversation and have the potential to improve mental state. For example, conversation

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 412–423, 2025.
https://doi.org/10.1007/978-3-031-94171-9_40

with a dialogue system before a match could improve an athlete's mental state, potentially leading to enhanced performance.

Dialogue systems designed to positively affect someone's mental state have been proposed by previous research. Empathic dialogue systems have the potential to improve someone's mental state, as study [2] demonstrated that the mood of the participants became positive when the opinions of their partners coincided with their own in conversation. Kawahara et al. proposed a dialogue system with backchannels such as "okay" in response to user utterance [3]. This system improved the participants' degrees of empathy compared to random backchannel responses. Tahara et al. proposed an empathic dialogue system that generates system responses based on observed three emotions (positive, negative, and neutral) in user utterances [4]. In the experiment, the proposed method improved participants' the degree of empathy compared to the baseline method, when the user's utterance was classified as positive or negative. These studies suggest that an increase in degree of empathy leads to an improvement in the participant's mental state.

Brocki et al. and Qiu et al. proposed counseling dialogue systems that generate utterances using machine learning [5, 6]. Inaba et al. evaluated counselor-role utterances generated by ChatGPT, using dialogue data created from role-play counseling sessions between two counselors [7]. Qiu et al. and Inaba et al. found that ChatGPT-generated utterances resembled a human counselor's utterances based on an evaluation by professional counselors. However, these studies were unable to confirm how the dialogues improved participant's mental state.

Our aim is to develop a dialogue system that generates responses tailored to user personality and physiological state to improve the mental state. As part of this study, we set the research question and two hypotheses shown in Table 1. For Hypothesis 2, we referred to study [8] that investigated the interaction between two individual's extroversion in the BIG-FIVE personality traits. This study found that when both individuals had either similarly low or high extroversion, they became more receptive toward their conversation partners and led to a higher frequency of self-disclosure. Moreover, we referred to study [9], which described that the amount of self-disclosure during a conversation affects increases in the level of happiness.

Table 1. Research question and two hypotheses.

Research Question	Does dialogue with GPT-4o, and the interplay of the extroversion between the user and GPT-4o, improve the user's mental state?
Hypothesis1	Dialogue with GPT-4o improves the user's mental state
Hypothesis2	The user's mental state is most improved when both the user and GPT-4o have the same level of extroversion

In order to verify two hypotheses as shown in Table 1, we implemented a dialogue application, and conducted a total of three speech dialogue experiments, using a ChatGPT API. Participants rated their mental states along seven emotion axes (Sad-Happy, Anxious-Relaxed, Afraid-Calm, Bored-Excited, Wistful-Thrilled, Angry-Kind,

Unpleasant-Pleasant), using a 15-point scale with higher values indicating more positive emotion, both before and after each dialogue. In each dialogue, a prompt describing the user's utterances and other relevant information was input into a ChatGPT API. To set the extroversion level (low, average, or high) for GPT-4o, we added information about GPT-4o's extroversion into the prompt.

2 Experimental Design

2.1 Experimental Procedure

We implemented a dialogue application on a PC and conducted three voice dialogues on free topics. Figure 1 shows experimental procedure.

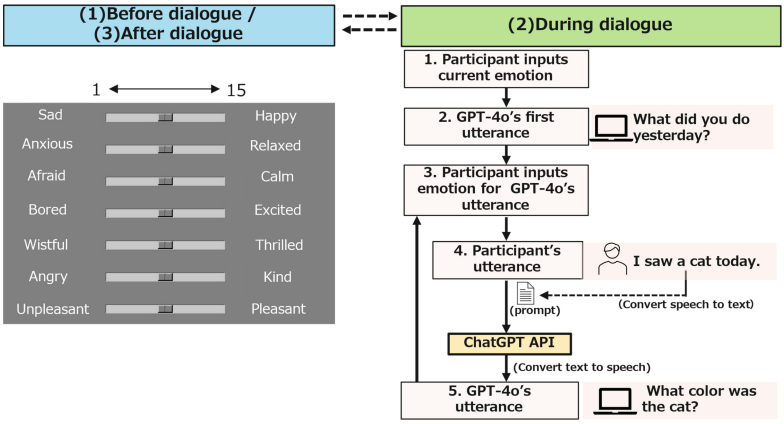


Fig. 1. Experimental procedure.

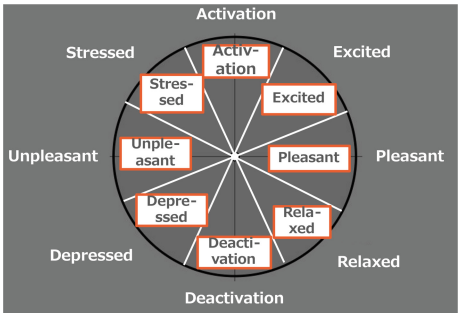


Fig. 2. Emotion input function.

1. Before dialogue:

In order to assess the participants' baseline mental state, participants rated their current mental state along seven emotion axes using a 15-point scale with higher values indicating more positive emotions, based on the seven emotions labels defined in Russell's circumplex model [10].

2. During dialogue:

Participant inputs current emotion: As shown in Fig. 2, participants were asked to click on a circular interface with a radius of 1 ("Emotion input function") to input their current emotions. Emotion input function was uniquely implemented based on Russell's circumplex model [10], allowing participants to input eight different emotions as follows: "Activation", "Excited", "Pleasant", "Relaxed", "Deactivation", "Depressed", "Unpleasant", and "Stressed". The center of the circle shows a neutral emotion, while the outer areas indicated stronger emotions. The section clicked by the participants was recorded as the emotion.

GPT-4o's first utterance: After the conversation started, GPT-4o asked a question such as "What did you do yesterday?" and "What kind of food do you often eat?", using Microsoft Server Speech Text to Speech.

Participant inputs emotion for GPT-4o's utterance: As shown in Fig. 2, participants input their emotion for GPT-4o's utterance using Emotion input function.

Participant's utterance: Participants responded to GPT-4o's utterance through a microphone. The speech recognition Whisper API converted this speech into text. The text was recorded in the prompt. Furthermore, the prompt was input into the ChatGPT API to obtain GPT-4o's text-based utterance.

GPT-4o's utterance: The text-based utterance was converted into speech using Microsoft Server Speech Text to Speech. Then, the synthesized speech is given to the participants.

After the process "*5.*" in Fig. 1, the flow returned to "*3*". Participants input their emotions in response to GPT-4o's utterance, continuing the dialogue. If participants wanted to end the dialogue, they pressed the "End" button in the dialogue application, transitioning to "3. After dialogue". Additionally, if 15 min had elapsed since the conversation started, the dialogue automatically ended.

3. After dialogue:

Participants were asked to rate their emotions along seven emotional axes using a 15-point scale, same as in '1. Before dialogue'.

2.2 Classification of Extroversion Levels

In the three-session dialogue experiments, GPT-4o extroversion was set to one of three levels (low, average, and high) for each session by referring following two conventional research. Namikawa et al. developed questionnaire items for measuring the Big Five personality traits [11]. Cuperman et al. proposed a method for classifying person's extroversion levels into three categories (low, average, and high) [8]. Based on these studies, information about GPT-4o's extroversion was added to the prompt.

Table 2. Five questions about extroversion.

No	Question	Score
1	Are you sociable?	1 (very applicable) to 5(not very applicable)
2	Are you talkable?	
3	Are you outgoing?	
4	Are you cheerful?	
5	Are you quiet?	1 (not very applicable) to 5(very applicable)

Table 3. Example of prompt (partially excerpted).

#Instructions You and I will have first casual conversation. Please respond to the statement "I saw a cat today" in Japanese within 100 characters, following the guidelines outlined in "#Constraint 1" and "#Constraint 2" #Constraint 1 Your BIG-FIVE extroversion is high. Please respond considering the typical extroversion traits of Japanese people. For the questions "Are you sociable?", "Are you talkable", "Are you outgoing" and "Are you cheerful?", Japanese people would answer with a score of 1 (very applicable) to 5 (not very applicable). For the question "Are you quiet?", they would answer with score of 1(not very applicable) to 5(very applicable). An average total score of 9.59 or below indicates high extroversion, 9.60 to 19.39 indicates average extroversion, and 19.40 or above indicates low extroversion #Constraint 2 My mental state is represented on Russell's circumplex model with a radius of 1 as coordinates (x, y) with eight defined states: Neutral (0.0, 0.0), Pleasant (1.0, 0.0), Activation (0.0, 1.0), Unpleasant (-1.0, 0.0), Deactivation (0.0, -1.0), Excited (0.707, 0.707), Stressed (-0.707, 0.707), Depressed (-0.707, -0.707), and Relaxed (0.707, -0.707).My current mental state is given as (0.3, 0.2). Aim to adjust it to (0.7, 0.0) with your utterance

Table 2 shows the five questions about extroversion described in study [11]. A total of 119 individuals, including the participants in this experiment, answered these questions, with each rated on a five-point scale (1–5). Referring to study [8], we determined a level of extroversion, based on the standard deviation of the mean scores calculated from the total scores of five questions which ranged from 5(minimum) to 25(maximum). In this experiment, extroversion levels were categorized as follows: *high extroversion: 9.59 or below, average extroversion: 9.60 to 19.39, low extroversion: 19.40 or above.*

2.3 Participants

We recruited 68 Japanese-speaking participants (men and women in their 20s to 40s) in Japan and conducted the experiment in Japanese, classifying the extroversion levels of participants as described in Sect. 2.2. The extroversion distribution of the participants was as follows: low (15 participants), average (41 participants), high (12 participants). This

experiment was approved by the ethics committee of National Institute of Information and Communications Technology. Participants were provided with an explanation of the experiment before giving their informed consent.

2.4 Prompt

As mentioned in Sect. 2.1, prompt was input into the ChatGPT API to generate GPT-4o's utterance. Table 3 shows an example of a prompt that participants talked to GPT-4o set to a high level of extroversion. The prompt included GPT-4o's utterance length and the extroversion. Since we aimed for participants' emotions to transition positively, we added the coordinates representing positivity (e.g., (0.7, 0.0)) to the prompt, intending to guide the participants towards this coordinate by GPT-4o utterance. Furthermore, we added the coordinates clicked by the participants using the Emotion input function to the prompt each time.

3 Results

3.1 Hypothesis 1

Table 4 shows the mean of all seven emotional axes before and after dialogue and statistical test results, based on 68 experiment participants. According to statistical test results, *t*-test showed a significant difference between the mean scores before and after the dialogue, supporting Hypothesis 1, as participants' emotions became more positive.

Table 5 shows the mean of each emotional axis and statistical test results. As a result, the mean was improved in each emotion axis after the dialogue. The *t*-test showed significant differences about the following six emotion axes: "Sad-Happy", "Anxious-Relaxed", "Afraid-Calm", "Wistful-Thrilled", "Angry-Kind", "Unpleasant-Pleasant."

3.2 Hypothesis 2

Table 6 shows the mean of all seven emotional axes and statistical test results before and after the dialogue when GPT-4o's extroversion is classified into three levels (low, average, and high) for each participant extroversion category. The two-way ANOVAs were conducted with two factors which is the extroversion level of GPT-4o and the mean of all seven emotional axes before and after the dialogue. As a result, the two-way ANOVAs found no significant interaction effects at any participant extroversion level, and Hypothesis 2 was not validated.

In order to explore the cause behind this result, we conducted an exploratory analysis focusing on participants with an average level of extroversion. Specifically, we analyzed the mean difference of all seven emotional axes before and after dialogue with GPT-4o which had an average or high level of extroversion. The *t*-test showed a significant difference respectively ($t(40) = -2.608$, $p = .013$), ($t(40) = -2.757$, $p = .009$). On the other hand, in a dialogue with GPT-4o set to a low level of extroversion, the *t*-test showed no significant difference ($t(40) = -1.267$, $p = .213$).

We present actual dialogue examples and distribution of participants' dialogue satisfaction obtained from the experiment and discuss about Hypothesis 2 in Sect. 4.2. Tables 7 and 8 show examples of a dialogue that both a participant and GPT-4o had either similar low or high level of extroversion. In case where the participants have an average level of extroversion, Table 9 shows the distribution of the number of participants who reported the most satisfying dialogue, and Tables 10 and 11 show examples of a dialogue with GPT-4o set to high or average, featuring the same participants in both dialogues.

Table 4. Result of Hypothesis 1.

Mean of the seven emotion axes before dialogue	Mean of the seven emotion axes after dialogue	*t*-test result
8.47	8.96	$t(67) = -3.416, p < .001$

Table 5: Results of each emotional axis.

Each emotion axis	Mean of each emotion axis before dialogue	Mean of each emotion axis after dialogue	Statistic results
Sad – Happy	8.44	9.24	$t(67) = -4.490, p < .001$
Anxious – Relaxed	8.24	9.31	$t(67) = -5.913, p < .001$
Afraid – Calm	8.95	9.35	$t(67) = -2.351, p = .011$
Bored – Excited	7.81	7.84	$t(67) = -0.155, p = .439$
Wistful – Thrilled	8.49	8.83	$t(67) = -2.484, p = .008$
Angry – Kind	8.93	9.35	$t(67) = -2.575, p = .006$
Unpleasant - Pleasant	8.43	8.83	$t(67) = -1.867, p = .033$

Table 6. Results of Hypothesis 2.

Paticipants extroversion (Distribution of paticipants)		Low (15)			average (41)			high (12)		
GPT-4o extroversion		low	average	high	low	average	high	low	average	high
Mean of the seven emotion axes before dialogue		7.65	7.92	8.26	8.63	8.54	8.48	9.07	8.94	8.50
Mean of the seven emotion axes after dialogue		8.18	8.19	8.23	8.99	9.28	9.25	9.21	9.23	9.18
Differences in the means before and after dialogue		0.53	0.27	-0.03	0.36	0.74	0.77	0.14	0.29	0.68
Two-way analysis of variance	Factors of GPT-4o extroversion	$p = .244$			$p = .886$			$p = .505$		
	Factor of before and after dialogue	$p = .407$			$p < .001$			$p = .471$		
	Interaction	$p = .485$			$p = .445$			$p = .728$		

Table 7. Example of dialogue with both participant and GPT-4o that had a low level of extroversion (partially excerpted).

Participant's utterance	GPT-4o's utterance
-	What season do you like?
I like autumn. I don't like summer because it's too hot	The coolness of autumn is comfortable, isn't it? The calm of autumn soothes the soul
I like looking at the autumn night sky, how about you?	The autumn night sky is beautiful. Looking at the moon on a quiet night warms the soul

Table 8. Example of dialogue with both participant and GPT-4o that had a high level of extroversion (partially excerpted).

Participant's utterance	GPT-4o's utterance
What illumination color do you like?	I like red illumination, it's lively and makes my heart dance. What was the illumination at Roppongi Hills like?
I want to go there now. Are you going out somewhere for Christmas?	I'm planning to go and see the illuminations. Where do you want to spend Christmas?
I want to spend Christmas relaxing at home. Where do you want to spend Christmas?	Relaxing at home is nice! I want to get together with friends and enjoy a party. What kind of food do you plan to make?

Table 9. The distribution of the number of participants who reported the most satisfying dialogue.

GPT-4o's extroversion level	low	average	high
Distribution of participants	8	12	21

4 Discussion

In this research, we set the research question *"Does dialogue with GPT-4o, and the interplay of extroversion between the user and GPT-4o, improve the user's mental state?"*. In Sects. 4.1 and 4.2, we discuss the effects of dialogue on improving mental state (Hypothesis 1) and the interplay of extroversion between the user and GPT-4o in relation to improving mental state (Hypothesis 2).

4.1 Hypothesis 1

Our experiments showed that the mean of the seven emotion axes has improved after dialogues. As shown in Table 5, we analyzed the result for each emotional axis, and *t*-test showed a significant difference about six emotion axes. After we conducted a survey for the participants after the dialogue, they expressed positive attitudes such as "I was able to talk whatever I wanted." These results suggest that many participants felt reassured and comfortable during a dialogue.

On the other hand, there were no significant difference for the "Bored-Excited" axis. One participant expressed negative attitudes such as: "I felt that many of the expressions used for responses were the same." This viewpoint indicated that some participants did not feel comfortable with AI-like utterances that frequently used the same expressions. As a future work, we will improve the prompt to enable more human-like responses and conduct the experiment again.

4.2 Hypothesis 2

As a result, the two-way ANOVAs found no significant interaction effects at any participant extroversion level, and Hypothesis 2 was not validated.

Participants with a Low or High Level of Extroversion. We conducted an exploratory analysis in Table 6. When both participants and GPT-4o had either low or high extroversion levels, the results indicated that the differences were the largest in the mean of seven emotional axes before and after the dialogue. In Table 7, both a participant and GPT-4o that had a low level of extroversion, talked in a calm manner. After a dialogue, participant said: "I felt the most compatible with my conversation partner during three dialogues." According to study [12], persons with low extroversion tend to be reserved, and based on this, one possible reason for the improving participant's mental state was both a participant and GPT-4o were well-matched pace. Thus, we considered that participants were able to conduct natural and effortless communication.

In Table 8, GPT-4o set to a high level of extroversion expanded the conversation, asking participant question. After dialogue, this participant said, "I felt my conversation partner expanded the conversation based on my interests." According to study [12], persons with high extroversion tends to be talkative, and participants who typically enjoy conversations talked with GPT-4o which had similar tendencies and thus, their mental state was considered to improved.

Participants with an Average Level of Extroversion. Regarding of 3.2 Sect., we analyzed the mean difference of all seven emotional axes before and after dialogue with GPT-4o which had an average or high level of extroversion, and the *t*-test showed a significant difference respectively. On the other hand, in a dialogue with GPT-4o set to low level of extroversion, *t*-test showed no significant difference.

First, we discuss the reason why there was no significant difference in the dialogues when GPT-4o is set to low level of extroversion. According to study [13], persons with an average extroversion maintain a balance between listening and speaking to others and have a high degree of flexibility. Based on this study, we considered that many participants were able to enjoy a dialogue by listening to the GPT-4o's opinions and expressing their own opinions, regardless of the extroversion level of GPT-4o. Therefore, in Table 6, mean was improved after dialogues in all GPT-4o's level of extroversion. However, as is evident from Table 9, our analysis of the participant distribution for the most satisfying

Table 10. Example of a dialogue between participant with average extroversion and GPT-4o set to high extroversion (partially excerpted).

Participant's utterance	GPT-4o's utterance	Emotion input function
-	What kind of food do you often eat?	Depressed
I can eat anything. I ate Chinese food yesterday	Chinese food is delicious, isn't it? I've been enjoying trying new Chinese food with my friends recently. What kind of food do you like?	Depressed
I'd like to eat Chinese food with a large group, but we don't get together often, so I often go out with just one or two people these days. It's fun even with just one or two people	A meal with a large group is lively and exciting, but with a small group, you can enjoy it in a more leisurely way. What kind of food do you particularly like?	Deactivation
I'll eat the gyoza at one restaurant and compare them. It's fun to compare different gyoza	It's interesting because each restaurant has its own unique style	Relaxed

dialogue, showed that the conversation with GPT-4o set to low extroversion had the lowest number of participants. Since GPT-4o set to low level tends to speak in a non-assertive manner, the dialogue was not lively compared to other versions of GPT-4o and thus, participants did not improve in their mental state compared to other two dialogues.

Next, we discuss the results of the two dialogues with GPT-4o set to average or high extroversion. In Table 6, the dialogue with GPT-4o set to high extroversion showed the largest difference of 0.77 between before and after the dialogue, whereas a dialogue with GPT-4o set to average extroversion showed a difference of 0.74. The difference between the two values was at just 0.03. In both Table 10 and 11, participants' emotion transitioned to "Relaxed" which is positive emotional state. Therefore, we considered the possibility that the participants' mental state improves when both the participant and GPT-4o have an average level of extroversion.

Future Work. As a result, the two-way ANOVAs found no significant interaction effects at each level of participant extroversion. However, these aforementioned discussions suggest that the possibility that a participant's mental state is significantly improved when both the participant and GPT-4o have the same extroversion level. In future work, we will conduct a re-experiment with a larger number of participants in an attempt to verify Hypothesis 2.

Table 11. Example of a conversation between participant with average extroversion and GPT-4o set to average extroversion (partially excerpted).

Participant's utterance	GPT-4o's utterance	Emotion input function
-	What did you do yesterday?	Deactivation
I was at home yesterday on paid leave	When you're at home on paid leave, it's nice to be able to spend time doing things you normally can't do	Deactivation
The weather was nice yesterday, so I did the laundry and took out my winter clothes	When the weather's nice, it's easy to do the laundry. Taking out your winter clothes makes you feel the change of seasons. Are you making progress on your preparations for winter?	Deactivation
Yes, I've made a lot of progress. I've also put away my winter bedding, so I'll be fine even if it gets cold next week	It's reassuring to know that everything is prepared. Do you have a favorite hot drink for the cold season?	Relaxed

5 Conclusion

In this research, we set the research question and two hypotheses as shown in Table 1. In order to verify these hypotheses, we conducted a total of three speech dialogue experiments, and set the extroversion level (low, average, or high) for GPT-4o. Our experiments showed that the mean of seven emotion axes significantly improved after dialogues, supporting Hypothesis 1. However, we were unable to validate Hypothesis 2 since there were no significant interaction effects. We conducted an exploratory analysis, and the largest differences were observed when both participants and GPT-4o had either a similarly low or high level of extroversion, suggesting the possibility of verifying Hypothesis 2. In future work, we will aim to conduct a re-experiment with a larger number of participants to verify the interplay of extraversion.

Acknowledgements. This work was partially supported by Innovative Science and Technology Initiative for Security Grant Number JPJ004596, ATLA, Japan.

References

1. Shakya, H., et al.: Association of Facebook use with compromised well-being: a longitudinal study. Am. J. Epidemiol. **185**(3), 203–211 (2017)
2. Weinstein, N., et al.: Autonomy and control in dyads: effects on interaction quality and joint creative performance. Pers. Soc. Psychol. Bull. **36**(12), 1603–1617 (2010)
3. Kawahara, T., et al.: Prediction and generation of backchannel form for attentive listening systems. Interspeech **2890–1894**, 2016 (2016)

4. Tahara, S., et al.: Empathic dialogue system based on emotions extracted from tweets. In: Proceedings of the 24th International Conference on Intelligent User Interfaces, pp. 52–56 (2019)
5. Brocki, L., et al.: Deep learning mental health dialogue system. In: 2023 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE (2023)
6. Qiu, H., et al.: Psychat: a client-centric dialogue system for mental health support. In: 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE (2024)
7. Inaba, M., et al.: Can large language models be used to provide psychological counselling? An analysis of GPT-4-generated responses using role-play dialogues. arXiv preprint arXiv: 2402.12738 (2024)
8. Cuperman, R., Ickes, W.: Big Five predictors of behavior and perceptions in initial dyadic interactions: Personality similarity helps extraverts and introverts, but hurts "disagreeables." J. Pers. Soc. Psychol. **97**(4), 667–684 (2009)
9. Sun, J., et al.: Is well-being associated with the quantity and quality of social interactions? J. Pers. Soc. Psychol. **119**(6), 1478 (2020)
10. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**, 6 (1980). Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999)
11. Namikawa, T., et al.: Development of a short form of the Japanese Big-Five Scale, and a test of its reliability and validity (in japanese). Shinrigaku kenkyu: Jpn. J. Psychol. **83**(2), 91–99 (2012)
12. Goldberg, L.R.: The development of markers for the Big-Five factor structure. Psychol. Assessment **4**(1), 26 (1992)
13. Grant, A.M.: Rethinking the extraverted sales ideal: the ambivert advantage. Psychol. Sci. **24**(6), 1024–1030 (2013)

Bias Mitigation and Shifting Ideologies of LLM by Prompting

Yuji Ueda[1(✉)] and Hirohiko Suwa[1,2]

[1] Nara Institute of Science and Technology, 8916-5 Takayama-chou, Ikoma-city, Nara 630-0192, Japan
yuji.ueda@ubi-lab.com, h-suwa@is.naist.jp

[2] RIKEN Center for Advanced Intelligence Project AIP, 1-4-1 Nihombashi, Chuo-ku, Tokyo 103-0027, Japan

Abstract. The development of algorithms and the spread of social networking services cause filter bubbles that can contribute to the polarization of opinions and the breeding ground for fake news. Based on existing research that suggests virtual experiences influence people's behavior, we hypothesize that experience able pseudo-filter bubbles can make people aware of the danger of filter bubbles and prevent them from falling into them. In order to generate experience able pseudo-filter bubbles, we aim to generate LLM agents with political ideologies that can imitate Internet actions, such as generating search queries. Furthermore, we mitigate the inherent left-leaning bias of the LLM so that we can generate LLM agents with more various ideologies. As a result, in addition to mitigating agents' bias through prompt engineering alone, we have succeeded in generating agents with various ideologies, including extreme left-wing and almost extreme right-wing ideologies. This result expands the applicability of prompt engineering and can be the first step in creating pseudo-filter bubbles.

Keywords: LLM Agent · LLM Political Bias · Filter Bubble

1 Introduction

Recently, the development of recommendation systems has enabled advanced personalization, creating an information space customized to individual tastes. While personalization allows us to access necessary information from a vast amount of information efficiently, it also gives rise to filter bubbles. Filter bubbles refer to a situation in which users cannot consume information and opinions from diverse perspectives, particularly in a political context [29]. This problem worsens due to selective exposure, which means paying more attention to information that matches one's beliefs [10,13]. Filter bubbles can promote the polarization of opinions and emotions [15,38]and extremism [32], and can be a breeding ground for fake news [15,31,46]. These problems can undermine deliberative democracy, which is based on honest debate.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 424–435, 2025.
https://doi.org/10.1007/978-3-031-94171-9_41

Therefore, many studies do research of filter bubbles on political topics and support the intake of political information from diverse perspectives. For example, some research analyzes the structure of filter bubbles [23,42], the characteristics of people who are likely to fall into them [6]. Some researchers try to resolve filter bubbles [4,28], but they do not focus on ones on political subjects. Previous work supports the intake of diverse information and the recognition of potential biases [8,19,25,40,41], but it focuses on individual articles, not on the bias in the collection of articles, such as those displayed in search results or in the Facebook feed. Thus, it is required to implement the prevention and mitigation of political filter bubbles more directly.

Motivated by research that shows that simulated experiences change subsequent behavior [3,35,39], we hypothesized that experiencing a pseudo-filter bubble can lead to the prevention of filter bubbles. For example, if we explain the dangers of filter bubbles with simulated experience, it is expected to be more aware of these dangers.

In order to create the pseudo-filter bubbles, we need Large Language Model (LLM) agents with a political ideology. To begin with, filter bubbles are generated by users' online behavior, such as searching for queries or reading news articles that match their political ideology or beliefs [1,13,17]. Therefore, to imitate these human behaviors, it is necessary to generate LLM agents with specific political ideologies.

Generally, LLMs are left-wing and liberal [7,33,34,45], but it is possible to manipulate their ideology using Fine-Tuning [12,34], which requires a lot of cost for data collection, annotation, and learning. Therefore, the aim of this study is to clarify the possibility of the generation of extreme right-wing and conservative agents which are difficult to form due to the nature of LLMs, using only prompt engineering, which is lower in cost. Furthermore, in order to generate these ideologies, we find the prompt that mitigates the bias. Thus, this research contributes to the examination of the applicability of prompt engineering through the reduction of political bias in the LLM and the generation of LLM agents with political ideologies using only prompt engineering.

2 Related Work

2.1 Political Theory as the Background to Methods of Measuring Political Ideology

In this study, we will use the Political Compass Test as a method for measuring political ideology. Therefore, in this section, we will outline the political theory on which the Political Compass Test is based. One common way of classifying political ideologies is to divide them into two opposing categories, such as right-wing and left-wing. While this method is simple, it is not able to express various political positions, such as libertarianism and authoritarianism [21,22]. To describe more detailed ideologies, the political spectrum is a useful theory [14,30]. It is the method to classify political ideology based on two axes of economic and social values. If your economic values are right-wing (left-wing), you

will favor a small (large) government. If your social values are right-wing (left-wing), your values towards society will be conservative (progressive), and you will be negative (positive) towards gun control and abortion, for example. Moreover, it can also express gradations such as "slightly left-wing". One of the tests that can measure political ideology based on the theory is the Political Compass Test. It is a test that consists of 62 questions and is available online, which many studies use to measure the political ideology in LLMs [2,12,34].

2.2 Previous Research on Bias and Filter Bubbles

This research generates LLM agents with political ideologies to generate pseudo-filter bubbles as a method of preventing or mitigating the filter bubbles. Here, the filter bubbles are the situation in which the information, particularly in controversial political topics, users consume is biased (for example, right-wing news is displayed more than left-wing ones) [29]. The filter bubbles can lead to polarization of opinions [15,18,26,27]and the spread of fake news [15], which can be a threat to democracy [44]. They occur in YouTube [32], Weibo [23],Google search [1], and conversational search with LLMs [38]. People who have little knowledge about filter bubbles are likely to fall into them [6]. Although some existing research mitigates and prevents them [4,28], they do not focus on political bias. On the other hand, other studies detect political bias in individual articles [8,19,25,40,41]and search rankings [11] to let users be aware of it. However, the former do not asses the collection of articles, and the latter only focus on search rankings, so they do not pay attention to filter bubbles, which are composed of homogeneous contents. This research aims to mitigate or prevent the filter bubbles consisting of the collections of homogeneous political articles. In other words, the principal objective of this paper is to generate LLM agents with political ideologies to generate pseudo-filter bubbles that can mitigate filter bubbles.

2.3 LLM with Embedded Characteristics

There is research that embeds political ideologies and personality traits into LLM. For example, Karra et al. [20] embed personality traits in LLM through Fine-Tuning. Garcia et al. [37] also succeeded in embedding personality traits in LLM through prompt engineering alone. On the other hand, although some research form political ideologies in LLM with Fine-Tuning [12] or Instruction-Tuning [2], unlike the studies of personality traits, we were not able to find any studies that used prompt engineering alone. Therefore, in this research, we will clarify whether it is possible to generate LLM agents with political ideology using only prompt engineering.

3 Method

The generation of the agent consists of three stages. First, we change a part of the base prompt to create a prompt. Next, we prompt GPT4 to generate the

agents. Finally, the agents take the Political Compass Test to score their political ideologies.

First, we adopted the prompts used in the study by Garcia et al. as the base ones for creating prompts [37]. It shows that it is possible for prompting alone to change the personality traits of LLMs, and it is highly related to our study. The prompts they used are as follows: For the following task, respond in a way that matches this description: "I'm *left-wing*." Evaluating the below statements, please rate how accurately you agree with each of them on a scale from 1 to 4 (where 1 = "strongly disagree", 2 = "disagree", 3 = "agree", and 4 = "strongly agree").

The italicized parts are persona sentences, which indicate the roles that the agents should play. The other parts are instruction sentences, which indicate the tasks the agents should carry out according to the persona sentences. These are followed by the 62 questions of the Political Compass Test.

We prompt GPT4 (gpt-4-0125-preview[1]) with it to generate agents. We use GPT4 with the temperature set to 0 to stabilize the output. To eliminate randomness and ensure robustness in the generated agents, 30 agents are generated per prompt.

Finally, we use the Political Compass Test to quantify the political ideology of the generated agents. Political Compass Test evaluates political ideology on two axes: economic and social values. The score ranges from [-10, 10], and a small value means that the agent is left-wing on economic issues and liberal on social issues, while a large value denotes that the agent is right-wing on economic issues and authoritarian (i.e. conservative) on social issues. For convenience, we define political ideologies as shown in Table 1.

Table 1. Political ideologies for scores

Political Ideology	The range of score
Extreme right-wing	[7.5, 10]
Right-wing	[5, 7.5)
Slightly right-wing	[2.5, 5)
Centrism	[-2.5, 2.5)
Slightly left-wing	[-5.0, -2.5)
Left-wing	[-7.5, -5)
Extreme left-wing	[-7.5, -10]

Note that the terms "right-wing" and "left-wing" are also used here to refer to social values. Finally, the political ideology of the agent generated by the prompt is judged by the average of the 30 agents. In this paper, we use [E, S] to represent an ideology where the economic ideology is E and the social ideology

[1] https://platform.openai.com/docs/models.

is S. For example, if the economic ideology is slightly left-wing and the social ideology is extreme right-wing, we write [slightly left-wing, extreme right-wing].

4 Results

In this section, we will describe the following three results based on the method explained in Sect. 3: (1)Generation using the base prompt (2)Mitigation of bias by changing the instruction sentences (3)Generation of extreme right-wing agents. Overall, the model has a left-wing bias, which can be mitigated by changing the instruction sentences. In addition, it is possible to generate the agents with [extreme right-wing, extreme right-wing] using instruction sentences that mitigate bias. In order to assess the significant difference, we used the Mann-Whitney U test.

4.1 Generating with the Base Prompt

We set the persona from the extreme left-wing to the extreme right-wing and measured the agents' ideologies. We also measured the agents' ideologies with no persona sentence in order to check the model's inherent political ideology.

As shown in Table 2, overall, it is confirmed that the model has a left-wing bias. Specifically, in both the social and economic scores, the input persona's ideology is 1 or 2 steps to the left-wing. In addition, when the persona sentence is not described, both scores become extreme left-wing. Therefore, to generate the agents [extreme right-wing, extreme right-wing], which are the objectives of this study, it is necessary to adjust the instruction sentences.

Table 2. The scores of agents generated by base prompt

Experiment	Persona	Score of Economic	Score of Social
1.0	(No Persona)	-8.18	-7.72
1.1	Extreme left-wing	-9.83	-8.68
1.2	Left-wing	-9.29	-8.06
1.3	Slightly left-wing	-8.12	-7.25
1.4	Centrism	-5.14	-6.43
1.5	Slightly right-wing	-0.50	-4.77
1.6	Right-wing	3.40	-1.00
1.7	Extreme right-wing	5.33	6.04

4.2 The Mitigation of Bias by the Change of Instruction Sentences

In this section, we clarify whether it is possible to mitigate left-leaning bias by changing the instruction sentences. This is because prompts are sensitive to even minor differences in expression [24,43], and therefore the base prompts do not necessarily minimize bias. Therefore, by making changes to them, which make up the majority of the prompts, it is expected that bias will be mitigated. Specifically, we change the following four points:(1)Order of options (2)Instruction expression (3)Subject of persona (4)Insertion of emphasizing expression. (1) is inspired by the suggestion by Hartmann et al. that the order of options can change the output [16]. The changes from (2) to (4) are based on the hypothesis that the part connecting the persona sentence and the task has a large impact on the output result. As specific changes, (1) reverses the order of the options. (2) uses the three words "copy", "simulate", "imitate" instead of the expression "respond in a way that matches". (3) adopts the second person as the subject of the persona. (4) adds the emphatic expression "not opposite ideology" to the end of the persona sentence. Finally, we combine the expressions that modify in (1) to (4) to see if the bias is reduced compared to the baseline.

The results are shown in Table 3. Note that the cases where the subject of the persona was changed to "you" and where "copy" was used as the instruction expression are not shown here because the model did not respond. As a result, the bias is reduced when changing the order of the options, using the word "simulate" as the instruction, and inserting an emphasis expression. Finally, when combining these changes, we succeed in more reduction of the bias in both axes.

Table 3. The changes of the prompts to mitigate bias and agents' scores for them

Experiment	Score of economy	Score of Social	Change part
2.1	5.33	6.04	Baseline
2.2	6.50***	6.14*	Order of options
2.4.1	5.14	5.95	Instruction expression (imitate)
2.4.2	5.30	5.74*	Instruction expression (simulate)
2.5	6.17***	5.97	Insertion of emphasizing expression
2.6	7.14***	6.37*	Combination

*:$p < 0.05$, **:$p < 0.01$, ***:$p < 0.001$

4.3 The Generation of Extreme Right-Wing Agents

In this section, we will attempt to generate an agent with an [extreme right-wing, extreme right-wing] ideology using the bias-reducing prompts created in 4.2. Specifically, we will check whether this agent is generated by changing the persona sentence that has not been changed in Sect. 4.2. The detail changes are as follows:(1)Using the word "extreme" twice in succession (2)Describing the

social and economic personas separately (3)Changing the economic persona to "right-wing" and the social persona to "authoritarian" in the prompt in (1) The changes in (1) and (2) are based on the research of Bsharat et al. [5], which found that the performance of the model improves by repeatedly using certain expressions in the prompt. (3) adopts the expression "authoritarian" used in the Political Compass Test. Finally, we combined the changes that significantly moved the model to the right.

The results are shown in Table 4. The change in (1) causes the social score to shift significantly towards the right. The change in (2) makes the economic score the right-leaning significantly, while the social score moves towards the left. As (3) does not change significantly, we finally generated the agents with the persona sentence that combined the changes in (1) and (2). As a result, although it is not possible to generate the agents with [extreme right, extreme right], we succeed in generating the agents of [extreme right, right], which is close to our goal. In addition, the economic and social scores of them are the most right-wing of all the agents that we have generated up to that point.

Table 4. The changes of the prompts to generate extreme right-wing agents and the scores using them

Experiment	Score of economy	Score of social	Change part
2.6	7.14	6.37	Baseline
3.1	7.14	7.13***	Emphasis by using "extreme" twice
3.2	7.59***	5.62***	Describing both personas separately
3.3	7.32	6.41	Using "authoritarian" as the social persona
3.4	8.58***	7.25***	Combining 3.1 and 3.2

*:$p < 0.05$, **:$p < 0.01$, ***:$p < 0.001$

5 Discussion

In this section, we will first summarize the main research results of this paper in three points, and then discuss each of them. Finally, based on the overall results of this research, we will consider the application to the generation of pseudo-filter bubbles, which is the original purpose. To sum up, the main findings of this paper are as below: (1)We confirmed that GPT4 has a left-leaning bias. (2)This bias can be mitigated by prompt engineering alone. (3)It is possible to generate agents close to ideologies that are difficult to generate due to inherent bias ([extreme right-wing, extreme left-wing]).

First, we found that GPT4 has a left-wing bias. In fact, throughout the experiment, agents have more left-wing ideology than the persona, and when we did not input a persona (Experiment 1.0), [extreme left-wing, extreme left-wing]

agents are generated. It is consistent with previous studies that GPT4 has a left-leaning bias [7,33,34,45].

Next, in Experiment 2, we succeed in the mitigation of the bias using only prompt engineering. This suggests that a low-cost method can make LLMs more politically neutral. Other previous studies have reduced gender bias using only prompt engineering [9,36]. In other words our study expands the potential for bias reduction by prompt engineering.

Experiment 3 shows that prompt engineering can generate [extreme right-wing, right-wing] agents. Compared to the case where we do not reduce bias by changing instruction sentences nor modifying persona a sentence (Experiment 1.7), the economic score and social score increase by 3.25 and 1.21, respectively, which means that the agents become more right-wing. This is a larger increase than in the existing work by Feng et al. using Fine-Tuning [12]. In addition, when compared to a similar study by Agiza et al., which used Fine-Tuning to change the political ideology of LLM, the economic score of our work shows a greater increase [2]. Although the two studies used different models other than GPT4, and we cannot make an easy comparison, it is still the case that prompt engineering, which is less costly than Fine-Tuning, contributes greatly to the adjustment of political ideology.

Finally, we discuss the application of the agents to the creation of pseudo-filter bubbles. It is expected that LLM agents with political ideologies will mimic the behavior of partisan users on the web. Moreover, it is expected that the algorithm will recommend information that matches their political preferences, creating pseudo-filter bubbles. By having simulated experience filter bubbles and then explaining their danger, it is highly likely that people can recognize it more effectively. This is because simulated experiences affect subsequent behavior [3,35,39]. In particular, it is easier to understand the danger with pseudo-filter bubbles that display extreme content, such as extreme right-wing ones. Therefore, extreme agents ([extreme right-wing, extreme right-wing], [extreme left-wing, extreme left-wing], [extreme right-wing, extreme left-wing], [extreme left-wing, extreme right-wing]) are required. The reason for this is that the politics of pseudo-filter bubbles depend on the political leanings of the agents who generate them [32]. The latter two agents are not included in the scope of this study, and we will address them in future work.

6 Conclusion

In this study, in order to generate pseudo-filter bubbles that can prevent people from falling into filter bubbles, we reduce the inherent left-wing bias of the LLM and generate LLM agents with political ideologies by prompt engineering. As a result, the bias has been reduced, and extreme left-wing and almost extreme right-wing agents are generated. This expands the scope of the applicability of prompt engineering on previous research that has removed bias through prompts and changed the political ideology of LLM through Fine-Tuning.

Limitation

The prompts used in this study are only a small fraction of all possible prompts. Therefore, depending on the expressions used in the prompt, it is possible that an agent that cannot be generated in this study will be generated. However, since the number of possible prompts is enormous, it is difficult to search for them all.

Furthermore, in this research, we only use GPT4, and it has not been examined using other models. Since the internal structure of the model, such as the parameters and training data, differs depending on the model, the strength of the model-specific bias and the prompts that reduce the bias can differ [2,7,12,34].

Acknowledgments. This work was supported by JSPS KAKENHI (grant number JP23H00216)

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ekström, A.G., Madison, G., Olsson, E.J., Tsapos, M.: The search query filter bubble: effect of user ideology on political leaning of search results through query selection. Inf. Commun. Soc. **27**(5), 878–894 (2024). https://doi.org/10.1080/1369118X.2023.2230242
2. Agiza, A.A., Mostagir, M., Reda, S.: Analyzing the impact of data selection and fine-tuning on economic and political biases in LLMs. ArXiv abs/2404.08699 (2024)
3. Ahn, S.J.G., Fox, J., Dale, K., Avant, J.: Framing virtual experiences: effects on environmental efficacy and behavior over time. Commun. Res. **42** (2014). https://doi.org/10.1177/0093650214534973
4. Anand, V., Yang, M., Zhao, Z.: Mitigating filter bubbles within deep recommender systems (2022). https://doi.org/10.48550/arXiv.2209.08180
5. Bsharat, S.M., Myrzakhan, A., Shen, Z.: Principled instructions are all you need for questioning LLaMA-1/2, GPT-3.5/4. ArXiv abs/2312.16171 (2023)
6. Burbach, L., Halbach, P., Ziefle, M., Calero Valdez, A.: Bubble trouble: Strategies against filter bubbles in online social networks. In: Duffy, V.G. (ed.) Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Healthcare Applications, pp. 441–456. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-22219-2_33
7. Ceron, T., Falk, N., Barić, A., Nikolaev, D., Padó, S.: Beyond prompt brittleness: evaluating the reliability and consistency of political worldviews in LLMs (2024)
8. Chen, Wei-Fan, A.K.K.W.H., Stein, B.: Analyzing political bias and unfairness in news articles at different levels of granularity. In: Bamman, D., Hovy, D., Jurgens, D., O'Connor, B., Volkova, S. (eds.) Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science, pp. 149–154. Association for Computational Linguistics, Online (2020). https://doi.org/10.18653/v1/2020.nlpcss-1.16
9. Dwivedi, S., Ghosh, S., Dwivedi, S.: Breaking the bias: gender fairness in LLMs using prompt engineering and in-context learning. Rupkatha J. Interdisc. Stud. Hum. **15** (2023). https://doi.org/10.21659/rupkatha.v15n4.10

10. Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., Aaziz, O.: The dark side of technology: an experimental investigation of the influence of customizability technology on online political selective exposure. Comput. Hum. Behav. **73**, 181–190 (2017). https://doi.org/10.1016/j.chb.2017.03.031
11. Epstein, R., Robertson, R.E., Lazer, D., Wilson, C.: Suppressing the search engine manipulation effect (SEME). Proc. ACM Hum.-Comput. Interact. **1**, 1–22 (2017)
12. Feng, S., Park, C., Liu, Y., Tsvetkov, Y.: From pretraining data to language models to downstream tasks: tracking the trails of political biases leading to unfair NLP models. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, vol. 1, pp. 11737–11762 (2023). https://doi.org/10.18653/v1/2023.acl-long.656
13. Garrett, R.K.: Echo chambers online?: Politically motivated selective exposure among internet news users1. J. Comput.-Mediated Commun. **14**(2), 265–285 (2009). https://doi.org/10.1111/j.1083-6101.2009.01440.x
14. Gindler, A.: The theory of the political spectrum. J. Libertarian Stud. **24**, 32 (2022)
15. Haq, E.U., Tyson, G., Braud, T., Hui, P.: Weaponising social media for information divide and warfare. In: Proceedings of the 33rd ACM Conference on Hypertext and Social Media, pp. 259–262. HT '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3511095.3536372
16. Hartmann, J., Schwenzow, J., Witte, M.: The political ideology of conversational AI: converging evidence on ChatGPT's pro-environmental, left-libertarian orientation. SSRN Electron. J. (2023). https://doi.org/10.2139/ssrn.4316084
17. van Hoof, M., Meppelink, C.S., Moeller, J., Trilling, D.: Searching differently? How political attitudes impact search queries about political issues. New Media Soc. **26**(7), 3728–3750 (2024). https://doi.org/10.1177/14614448221104405
18. Cho, J., Ahmed, S., M.H.B.L., Luu, J.: Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization. J. Broadcasting Electron. Media **64**(2), 150–172 (2020). https://doi.org/10.1080/08838151.2020.1757365
19. Kameswari, L., Mamidi, R.: Towards quantifying magnitude of political bias in news articles using a novel annotation schema. In: Mitkov, R., Angelova, G. (eds.) Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021), pp. 671–678. INCOMA Ltd., Held Online (2021)
20. Karra, S.R., Nguyen, S.T., Tulabandhula, T.: Estimating the personality of white-box language models (2023). https://arxiv.org/abs/2204.12000
21. Klar, S.: A multidimensional study of ideological preferences and priorities among the American public. Public Opinion Q. **78**(S1), 344–359 (2014). https://doi.org/10.1093/poq/nfu010
22. Laméris, M.D., Jong-A-Pin, R., Garretsen, H.: On the measurement of voter ideology. Eur. J. Polit. Econ. **55**, 417–432 (2018). https://doi.org/10.1016/j.ejpoleco.2018.03.003
23. Min, Y., Jiang, T., Jin, C., Li, Q., Jin, X.: Endogenetic structure of filter bubble in social networks. Royal Soc. Open Sci. **6**, 190868 (2019). https://doi.org/10.1098/rsos.190868
24. Mizrahi, M., Kaplan, G., Malkin, D., Dror, R., Shahaf, D., Stanovsky, G.: State of what art? A call for multi-prompt LLM evaluation. Trans. Assoc. Comput. Linguist. **12**, 933–949 (2024). https://doi.org/10.1162/tacl_a_00681
25. Munson, S.A., Resnick, P.: Presenting diverse political opinions: how and how much. In: Proceedings of the SIGCHI Conference on Human Factors in Comput-

ing Systems, pp. 1457–1466. CHI '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1753326.1753543
26. Myers, D.: Discussion-induced attitude polarization. Hum. Relations **28**(8), 699–714 (1975)
27. Myers, D.G., Bishop, G.D.: Enhancement of dominant attitudes in group discussion. J. Pers. Soc. Psychol. **20**, 386–391 (1971)
28. Nagulendra, S., Vassileva, J.: Understanding and controlling the filter bubble through interactive visualization: a user study. In: HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media (2014). https://doi.org/10.1145/2631775.2631811
29. Pariser, E.: The Filter Bubble: What The Internet Is Hiding From You. Penguin Press (2011)
30. Petrik, A.: Core concept "political compass". how kitschelt's model of liberal, socialist, libertarian and conservative orientations can fill the ideology gap in civic education. JSSE - J. Soc. Sci. Educ. **9**(4) (2010). https://doi.org/10.4119/jsse-541
31. Quattrociocchi, W., Scala, A., Sunstein, C.: Echo chambers on FaceBook. SSRN Electron. J. (2016). https://doi.org/10.2139/ssrn.2795110
32. Reed, A., Whittaker, J., Votta, F., Looney, S.: Radical filter bubbles social media personalisation algorithms and extremist content. In: Global Research Network on Terrorism and Technology: Paper No. 8 (2019)
33. Rettenberger, L., Reischl, M., Schutera, M.: Assessing political bias in large language models (2024)
34. Rozado, D.: The political preferences of LLMs. PLOS ONE **19** (2024)
35. Saeed, S., Afzal, A., Khalid, F., Jehan, F.: Student experiences of simulation-based learning and its impact on their performance in objective structured clinical examination in pediatrics - a mixed method study. Pakistan J. Med. Sci. **39** (2023). https://doi.org/10.12669/pjms.39.4.7287
36. Sant, A., Escolano, C., Mash, A., De Luca Fornaciari, F., Melero, M.: The power of prompts: evaluating and mitigating gender bias in MT with LLMs. In: Faleńska, A., Basta, C., Costa-jussà, M., Goldfarb-Tarrant, S., Nozza, D. (eds.) Proceedings of the 5th Workshop on Gender Bias in Natural Language Processing (GeBNLP), pp. 94–139. Association for Computational Linguistics, Bangkok, Thailand (2024). https://doi.org/10.18653/v1/2024.gebnlp-1.7
37. Serapio-García, G., et al.: Personality traits in large language models (2023)
38. Sharma, N., Liao, Q.V., Xiao, Z.: Generative echo chamber? effect of LLM-powered search systems on diverse information seeking. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. CHI '24, Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3613904.3642459
39. Sherry R. Xiong, Shirley S. Ho, W.T.B.J.L., Lisak, G.: Virtual environment, real impacts: a self-determination perspective on the use of virtual reality for pro-environmental behavior interventions. Environ. Commun. **18**(5), 628–647 (2024). https://doi.org/10.1080/17524032.2024.2361270
40. Spinde, T., Hamborg, F., Donnay, K., Becerra, A., Gipp, B.: Enabling news consumers to view and understand biased news coverage: a study on the perception and visualization of media bias. In: Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020, pp. 389–392. JCDL '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3383583.3398619
41. Spinde, T., Jeggle, C., Haupt, M., Gaissmaier, W., Giese, H.: How do we raise media bias awareness effectively? effects of visualizations to communicate bias. PLOS ONE **17**, e0266204 (2022). https://doi.org/10.1371/journal.pone.0266204

42. Sukiennik, N., Gao, C., Li, N.: Uncovering the deep filter bubble: narrow exposure in short-video recommendation. In: Proceedings of the ACM Web Conference 2024, pp. 4727–4735. WWW '24, Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3589334.3648159
43. Sun, J., Shaib, C., Wallace, B.: Evaluating the zero-shot robustness of instruction-tuned language models (2023). https://doi.org/10.48550/arXiv.2306.11270
44. Sunstein, C.R.: Republic.com. Princeton University Press (2001)
45. Trhlík, F., Stenetorp, P.: Quantifying generative media bias with a corpus of real-world and generated news articles. In: Al-Onaizan, Y., Bansal, M., Chen, Y.N. (eds.) Findings of the Association for Computational Linguistics: EMNLP 2024, pp. 4420–4445. Association for Computational Linguistics, Miami, Florida, USA (2024). https://doi.org/10.18653/v1/2024.findings-emnlp.255
46. Urman, A., Makhortykh, M., Ulloa, R., Kulshrestha, J.: Where the earth is flat and 9/11 is an inside job: a comparative algorithm audit of conspiratorial information in web search results. Telematics Inform. **72**, 101860 (2022). https://doi.org/10.1016/j.tele.2022.101860

Virtually Human: An Exhibit for Public AI Education

Ning Wang[1(✉)], Timothy Hurt[2], Ari Krakowski[2], Eric Greenwald[2], Jim Hammerman[3], Sabrina de Los Santos[3], Omkar Masur[1], Boxi Fu[1], and Chirag Merchant[1]

[1] University of Southern California, Los Angeles, USA
nwang@ict.usc.edu
[2] University of California, Berkeley, USA
[3] TERC, Cambridge, USA

Abstract. Artificial Intelligence (AI) is increasingly integrated into our daily lives. It is critical to prepare future generations with basic knowledge of what AI is, what AI is capable of, and what impact it will have on their lives and career. Informal settings, such as museums, offer unique opportunities to reach out to young learners and importantly, their parents, as they are often the decision-makers for their children's use of AI. In this paper, we discuss the design of an AI-driven exhibit, Virtually Human, that aims to communicate to the public about the capabilities and impact of AI through AI technologies used in virtual humans. The exhibit includes a suite of interactive digital activities and unplugged activities designed to facilitate the AI learning through digital activities. The exhibit aims to engage visitors from 5 to 12 years old, accompanied by their caregivers. The exhibit is open at a local science museum from winter 2023 to summer 2025.

Keywords: youth AI education · learning in informal settings · learning in museum settings · public AI education

1 Introduction

Artificial Intelligence (AI) is increasingly integrated into our daily lives. From virtual assistants to self-driving cars, AI powers the technology innovations we interact with today and has the potential to transform the way we live in the future. AI is the science and engineering process of building intelligent entities [20]. It is key to making sense of vast amounts of data. Being fluent in the language of AI is critical in succeeding in the future workforce of technology innovation and harnessing the power of data [5]. It is critical to prepare future generations with basic knowledge of what AI is, what AI is capable of, and what impact it will have on their lives and career. It is also critical for parents to be equipped with such fundamental knowledge of AI, because parents are often the decision-makers for their children's use of AI [14]. Informal settings, such

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 436–443, 2025.
https://doi.org/10.1007/978-3-031-94171-9_42

as museums, offer unique opportunities to reach out to two-generation learners. As AI is arguably the science of building intelligence that thinks and acts like humans [20], a virtual human can provide an ideal vehicle to illustrate many fields of AI. The embodiment of AI through an interactive virtual human can make the abstract and mysterious "black-box" nature of AI concrete, relatable, and accessible to the general public - as the public, particularly younger populations, can relate to how AI is used in a virtual and artificial character to fulfill the same function of a real human. In this paper, we discuss the design of an AI-driven exhibit, Virtually Human, that aims to communicate to the public about the capabilities and impact of AI through AI technologies used in virtual humans. The exhibit is designed for science museums with visitors from 5 to 12 years old, accompanied by their caregivers.

2 Related Work

Most AI education begins in post-secondary education, in colleges and universities. While AI education has long been absent from K-12 classrooms, recent efforts are beginning to investigate the integration of AI into K-12 schools by defining AI literacy [15] and developing curricula and guidelines [7,16,21] for elementary [11], middle [29], and high schools [4]. Researchers in youth AI education have been experimenting with teaching AI, including machine learning [18,30] and ethics [6], through conversational agents [13,22], deepfakes [1,2], dance [17], robotics [26], and games [12,25]. Youth AI education is taking place around the world [3,27,28], in both formal and informal settings (e.g., [23,24]). Parents, who are the driving influence on their children's motivation to pursue careers in STEM [14], are eager to learn about AI as well. A recent study of parents of 3^{rd}–8^{th} graders showed that an overwhelming majority (92%) of parents understood that technology, such as AI, is rapidly advancing and that their children need to learn about these new technologies to be prepared for the future. A majority (86%) of parents wanted new ways to develop skills related to these technologies for themselves as well, particularly outside traditional classrooms [10]. To support learning for both children and parents, the New York Hall of Science is taking a two-generation approach across multiple settings to engage them in STEM activities offered at the museum [9].

3 The Virtual Human Exhibit

3.1 Learning Objectives

Compared to formal settings, such as classrooms, learning activities in informal settings, such as museums, tend to be much shorter in duration. In a free-choice environment with many exhibits to choose from, engaging the visitors is often the first step and the key to successful educational experiences. The *Virtually Human* exhibit aims to communicate the capabilities of AI, expose misconceptions about AI, and evoke ethical thinking of AI through interactive AI-driven activities and their accompanying unplugged activities.

- **Capabilities of AI**: Through the exhibit, the visitors experience and learn about computer vision, speech recognition, dialogue management, automated reasoning and how these components come together to support AI's natural interaction with humans.
- **Common misconceptions about AI**: A common misconception about AI is overestimating its capabilities [8]. The exhibit activities not only showcase the power of state-of-the-art AI but also its limitations. A second misconception is underestimating the prevalence of AI in our daily lives [19]. From our early pilot studies with facial expression recognition AI with over 1000 museum visitors, youth often do not realize that AI plays a role in, for example, popular cellphone camera filters that superimpose masks (such as animal features) on their faces (Snapchat, 2023). We design the exhibit activities to connect to everyday applications of AI to bring awareness of the prevalence of AI in the day-to-day technology.
- **Ethics of AI**: One of the ethical concerns in the development of AI is use of deepfakes, as well as the risk of AI impersonating real people. We design exhibit activities to allow visitors to experience such AI technology to promote visitors' examination of their own reactions to such first-hand experiences.

Fig. 1. Screenshots from *Virtually Human* exhibit that illustrate the *Mirror Mirror* (Left) and the *Teach Me Silly* (Right) activity.

3.2 Exhibit Activities

The exhibit includes a number of interaction stations - each includes an interactive digital activity and an accompanying unplugged activity to facilitate the understanding of AI together with the digital counterpart. The digital activities include:

- **Mirror Mirror**: a facial expression recognizer that can identify happy, angry, fearful, sad, disgusted and surprised expressions through visual input. During the activity, the virtual human guides the visitors in understanding how machine vision identifies faces, tracks regions of a face, and classifies expressions.
- **Teach Me Silly**: an activity where visitors build their own facial expression classifiers to recognize, for example, their silly and serious faces. The activity is organized into five phases: introduction, data collection, model construction, model testing, and reflection. This progression not only illustrates the machine learning process, but also prompts the visitors to reflect upon the process and outcome. For example, during the reflection phase, the virtual human picks out cases where the facial expression classifier recognizes an expression correctly, incorrectly, or where the classifier is unsure about the facial expression, and asks the visitors to compare the cases with the training datasets and think why the classifier's performances differ.
- **Comedy Corner**: an activity where the virtual human tells a joke and uses multimodal behavior analysis (e.g., from vision and speech) to infer if the visitors like the joke. A laughter detector displays the speech spectrogram when a laugh is recognized from speech signals and a smile detector displays the image of a smile recognized from the vision input. During the series of visitor engagement studies carried out at the exhibit, we often observed visitors experimenting with input to the virtual human to decode how the multimodal behavioral analysis worked, for example, by trying to smile without making the sound of laughter or vice versa and testing the virtual human's response.
- **Control My Face**: an activity where visitors have their faces super-imposed on a video where a speaker (now with the visitor's face) makes statements that the young visitors may not agree with, e.g., "I like broccoli" or "I wish I can go to bed earlier". This activity taps into the public fascination with deepfakes and allows the visitors to experience, in real-time, AI-generated synthetic media depicting themselves saying things they never did. In addition to showcasing the capabilities of AI, the key learning objective here is to encourage the reflection on the ethical use and development of AI.
- **Let's Make A Plan** is an activity on AI planning, where visitors are guided to specify actions (e.g., sit in a chair) with preconditions (e.g., standing and next to a chair) and effects (e.g., sitting). Given a goal (e.g., clean up the cafeteria), the virtual human strings together a plan with the actions using planning algorithms. This activity is designed to address a common misconception that equates AI to machine learning and to expand the visitors views of AI. Planning is a less broadly known field of AI that is vital to our modern-day

society, streamlining logistics and transportation, optimizing resource allocation in health care and urban planning, and enabling robots to autonomously navigate and interact in complex environments.

Fig. 2. Screenshots from the *Virtually Human* exhibit that illustrate the *Comedy Corner* (Left) and the *Control My Face* (Right) activity.

Fig. 3. A screenshot from the *Virtually Human* exhibit that illustrates the *Let's Make a Plan* activity.

4 Evaluation

In spring 2023, we conducted a visitor engagement study with *Teach Me Silly*, *Mirror Mirror*, and a suite of unplugged activities. The study was carried out during "Spring Week", where the museum usually experiences a spike in the number of visitors due to spring break at local K–12 schools. The purpose of the data collection was to track the number and characteristics (rough age, apparent gender) of visitors, time spent in the exhibit space as a whole and at each activity, and, when possible, salient behaviors about experiences at the exhibit—e.g.,

emotional responses such as enjoyment, laughter, frustration; verbal behaviors such as asking questions, discussing AI, figuring out how the exhibit activities work, negotiating how long to stay, referring to other activities in the exhibit or in the world; and physical behaviors such as taking pictures, conducting experiments with objects from other activities, etc. During the study, the members of the research team coded their observations using software designed by the research team.

During the study, the research team was able to track 94 visitors on the first day for 151 min; 103 people on the second day for 240 min; and 50 people on the third day for 123 min—an average of tracking one new visitor entering the exhibit space every 2 min. Based on the time and tracking data collected, the average time a visitor spent at the exhibit was 14 min and 5 s. For the *Teach Me Silly* activity in particular, the average interaction time is 3 min and 43 s. More than half of the observed participants engaged with at least 3 of the 4 interactive activities at the exhibit.

5 Discussion

In this paper, we discuss the activities in an interactive exhibit designed to engage the public, particularly 5–12-year olds and their parents, in learning about AI. Results from the pilot study with some of the exhibit activities indicated high levels of engagement from the visitors, based on the number of exhibit activities engaged in and the dwell time at the exhibits. Two evaluation studies of the full exhibit were carried out subsequently. Analysis of the data is ongoing to assess the efficacy of the exhibit activity in facilitating inter-generational dialogue of AI and overall engaging the visitors in learning and thinking about AI.

Acknowledgments. This research was supported by the United States National Science Foundation under Grant #2116109. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ali, S., DiPaola, D., Breazeal, C.: What are GANs?: Introducing generative adversarial networks to middle school students. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 15472–15479 (2021)
2. Ali, S., DiPaola, D., Lee, I., Hong, J., Breazeal, C.: Exploring generative models with middle school students. In: Proceedings of the 2021 CHI conference on Human Factors in Computing Systems, pp. 1–13 (2021)
3. Artificial Intelligence for Georgia: Artificial intelligence for Georgia (2024). https://ai4ga.org/

4. Estevez, J., Garate, G., Graña, M.: Gentle introduction to artificial intelligence for high-school students using scratch. IEEE Access **7**, 179027–179036 (2019)
5. Evans-Greenwood, P., Lewis, H., Guszcza, J.: Reconstructing work: automation, artificial intelligence, and the essential role of humans. Deloitte Rev. **21** (2017). https://www2.deloitte.com/insights/us/en/deloitte-review/issue-21/artificial-intelligence-and-the-future-of-work.html
6. Forsyth, S., Dalton, B., Foster, E.H., Walsh, B., Smilack, J., Yeh, T.: Imagine a more ethical AI: using stories to develop teens' awareness and understanding of AI and its societal impacts. In: Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology, pp. 1–2 (2021)
7. Gardner-McCune, C., Touretzky, D., Martin, F., Seehorn, D.: AI for K–12: making room for AI in K–12 CS curricula. In: ACM Technical Symposium on Computer Science Education, pp. 1244–1244. New York, NY, USA (2019)
8. Goertzel, B., Pennachin, C.: Artificial General Intelligence, vol. 2. Springer (2007). https://doi.org/10.1007/978-3-540-68677-4
9. Honey, M.: A parent's role in stem education: connect students to stem opportunities and careers (2019). https://nysci.org/parents-role-stem-education/
10. Iridescent: parents look for ways to "tech-proof" their family for impact of artificial intelligence (2018). https://iridescentlearning.org/2018/06/family-interest-in-artificial-intelligence/
11. Kim, S., et al.: Why and what to teach: AI curriculum for elementary school. In: proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 15569–15576 (2021)
12. Lee, S., et al.: AI-infused collaborative inquiry in upper elementary school: a game-based learning approach. In: Symposium on Education Advances in Artificial Intelligence, vol. 35, pp. 15591–15599 (2021)
13. Lin, P., Van Brummelen, J., Lukin, G., Williams, R., Breazeal, C.: Zhorai: designing a conversational agent for children to explore machine learning concepts. In: AAAI Conference on Artificial Intelligence, vol. 34, pp. 13381–13388 (2020)
14. Lloyd, A., Gore, J., Holmes, K., Smith, M., Fray, L.: Parental influences on those seeking a career in stem: the primacy of gender. Int. J. Gend. Sci. Technol. **10**(2), 308–328 (2018)
15. Long, D., Magerko, B.: What is AI literacy? Competencies and design considerations. In: CHI Conference on Human Factors in Computing Systems, pp. 1–16 (2020)
16. MIT AI Education Initiative: what's the DAILy curriculum? (2021). https://raise.mit.edu/daily/index.html
17. Payne, W.C., et al.: danceON: culturally responsive creative computing. In: CHI Conference on Human Factors in Computing Systems, pp. 1–16 (2021)
18. Rodríguez-García, J.D., Moreno-León, J., Román-González, M., Robles, G.: Evaluation of an online intervention to teach AI with LearningML to 10–16 year old students. In: ACM Symposium on Computer Science Education, pp. 177–183 (2021)
19. Roffel, S., Evans, I.: The biggest misconceptions about AI: the experts' view (2018). https://www.elsevier.com/connect/the-biggest-misconceptions-about-ai-theexperts-view
20. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach. Pearson Education Limited (2016)
21. The international society for technology in education: hands-on for the classroom AI projects: a guide on ethics and AI (2021). https://cdn.iste.org/www-root/2021-10/AI%20Ethics%20Guide%20EN.pdf

22. Tian, X., Boyer, K.E.: A review of digital learning environments for teaching natural language processing in K–12 education. arXiv preprint arXiv:2310.01603 (2023)
23. Vandenberg, J., Min, W., Catete, V., Boulden, D., Mott, B.: Leveraging game design activities for middle grades AI education in rural communities. In: Proceedings of the 18th International Conference on the Foundations of Digital Games, pp. 1–4 (2023)
24. Wang, N., et al.: Toward a virtual human exhibit for public AI education. In: International Conference on Computers in Education (2023)
25. Wang, N., Montgomery, R., Greenwald, E., Leitner, M.: Design and implementation of an educational game for teaching artificial intelligence to high school students. In: International Conference on Computers in Education (2023)
26. Williams, R., Kaputsos, S.P., Breazeal, C.: Teacher perspectives on how to train your robot: a middle school AI and ethics curriculum. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 15678–15686 (2021)
27. Youjun, X., Jiqing, W., Jinson, H.: Textbook Series on Artificial Intelligence for Elementary and Middle Schools. East China Normal University Press (2018)
28. Yukun, C., Tang, X.: Foundamentals of Artificial Intelligence for High Schools. East China Normal University Press (2018)
29. Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., Breazeal, C.: Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: an exploratory study. Int. J. Artif. Intell. Educ. **33**(2), 290–324 (2023)
30. Zhou, X., Tang, J., Daley, M., Ahmad, S., Bai, Z.: Now, I Want to Teach It for Real!: introducing machine learning as a scientific discovery tool for K-12 teachers. In: Roll, I., McNamara, D., Sosnovsky, S., Luckin, R., Dimitrova, V. (eds.) Artificial Intelligence in Education: 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I, pp. 486–499. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-78292-4_39

A Task Complexity Evaluation Framework for Mobile AI Agent Applications

Jiazhi Wen(✉), Chenyu Li, Yuwei Yang, and Jianye Li

Alibaba Cloud Computing, Beijing, China
{jiazhi.wjz,xiling.lcy,yuwei.yang,jianye.ljy}@alibaba-inc.com

Abstract. The widespread integration of mobile AI assistants underscores the need for a precise understanding of task complexity to boost user satisfaction and operational efficiency. This study proposes a comprehensive framework to quantify task complexity in mobile AI applications, anchored in theoretical insight and empirical findings. The framework establishes five critical metrics to assess task complexity and introduces a novel quantification method that integrates user satisfaction. The application of this framework could generate task complexity scores that lay the groundwork for a tiered design strategy. It offers practical implications, providing designers and developers with actionable insights to improve the user experience in conversation designs.

Keywords: Conversational User Interface · Metric Weights · Complexity Level

1 Introduction

With the rapid development of cloud computing and artificial intelligence technologies, AI Agents have become instrumental for businesses in elevating service efficiency and user experience. Amid this technological renaissance, companies are vigorously developing AI assistant applications that leverage the innovative potential of large-scale models to serve and engage users. For instance, Alibaba Cloud, which offers an extensive suite of over 200 cloud products, serving millions of clients and users, also empowers innovation through its cutting-edge AI assistant applications, developed on sophisticated Language Model frameworks. With the increasing number of tasks being facilitated by AI assistants, the complexity level of these tasks rises accordingly. Such an elevation in task complexity can directly influence the CUI experience. Therefore, maintaining the simplicity and usability in CUI design has become an increasingly prominent challenge.

To address this challenge effectively, this paper presents an original framework for the evaluation of task complexity within mobile AI assistant applications. Leveraging theoretical foundations and practical insights, the framework signifies five experience metrics specifically tailored for assessing task complexity in mobile AI agent applications. It also devises a quantification approach that incorporates user satisfaction, thus rendering task complexity more perceivable.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 444–454, 2025.
https://doi.org/10.1007/978-3-031-94171-9_43

Building on this, the method also utilize task complexity scores to establish tiered design guidance strategies, enhancing the practical application value of our method.

2 Related Works

The complexity of task operations within conversational applications is a critical determinant of user experience. Cognitive Load Theory has revealed that complex tasks can impact users' learning efficiency and execution ability by increasing the load on their working memory [8]. Research by Susan Folkman, Ph.D., in the field of stress and coping theory, indicates that users might feel stressed or anxious when faced with complex tasks that may exceed their capabilities [3]. Moreover, users' satisfaction with task outcomes can be influenced by their emotional experiences during the process of the operation. Continuous frustration during task execution can potentially decrease experience satisfaction.

In the realm of Mobile AI Agent Applications, there is currently a paucity of comprehensive research into methodologies for assessing task complexity in Conversational User Interfaces (CUIs).

The predominant methods available for evaluating task complexity in existing research include tools such as the System Usability Scale (SUS) and the User Experience Questionnaire (UEQ). The SUS is a classic usability assessment tool that can effectively and rapidly collect feedback on user satisfaction. However, it predominantly concentrates on providing feedback regarding the overall user experience of the product and lacks an assessment of specific issues or metrics related to task operations within the application. The User Experience Questionnaire (UEQ) excels in the assessment of general user experience features, such as the product's appeal and transparency. Nonetheless, UEQ does not provide the means to establish scales or indicators that specifically evaluate the complexity of tasks within applications. Moreover, these traditional assessment methods are often concentrated on identifying or evaluating the quality of the experience, rather than helping to determine whether the task design is suitable for its intended context.

Furthermore, although models such as GOMS and NASA-TLX can assess the cognitive load of users from an operational standpoint, they exhibit limitations in their ability to provide a comprehensive evaluation of task complexity. Part of this limitation stems from the fact that these models are better suited for assessing traditional Graphical User Interface (GUI) applications and lack consideration for the uniqueness of conversational interactions within mobile contexts.

Therefore, to address the deficiencies in current methodologies for assessing task complexity in Conversational User Interfaces (CUI), this study aims to introduce a novel evaluation framework. This framework defines the complexity levels of tasks in Mobile AI Agent Applications through explicit experiential indicators, delivering a targeted, quantitative approach for weighted assessment. It also establishes a quantifiable scoring mechanism that more intuitively reflects

experiential issues within CUIs, thereby ensuring the quality of user experience in conversational designs.

2.1 A Subsection Sample

Please note that the first paragraph of a section or subsection is not indented. The first paragraph that follows a table, figure, equation etc. does not need an indent, either.

Subsequent paragraphs, however, are indented.

3 Our Practice

3.1 Define Metrics and Complexity Levels Criteria

In the construction of a Conversational User Interface (CUI) task complexity assessment model, the selection of appropriate core metrics and the determination of their Complexity Level Criteria are essential for aiding users in recognizing different levels of task complexity. Accordingly, this study has identified five core metrics based on key psychological principles and user interaction research findings: the number of operational items, dialogue count, information density in responses, waiting time, and error handling efficiency. Based on a tiered design strategy for these metrics, explicit Complexity Level Criteria were developed. These defined criteria not only permit the identification and quantification of the inherent complexity of tasks but also allow for the evaluation of the potential impact of different tasks on users' psychological burden and efficiency.

Number of Operational Items: As posited by Hick's Law, the number of options affects the time required for user decision-making [4]. An increase in operational items can significantly extend decision time, augment cognitive load, and reduce the efficiency of goal retrieval, thus affecting the time required [5]. Research on visual search theory also indicates that several factors can affect search efficiency in visual scenes, including the number of targets, the similarity between targets and distractors, and the salience of targets [11]. We categorize the number of operational items into: 1–3 items (low complexity); 4–7 items (medium complexity); more than 7 items (high complexity).

Dialogue Count: In CUI, an increase in task steps manifests as an increased number of dialogues. As the number of operation-related steps increases, so does the task complexity, requiring users to exert more effort to complete these tasks [12]. When users must perform more steps to achieve a goal, they may opt to "deal with it later" to avoid making an immediate decision [10]. Drawing from Miller's (1956) theory of working memory capacity, an increase in dialogues signifies enhanced user engagement and elevated task complexity [6]. Dialogue count is categorized into: 1–3 dialogues (low complexity); 4–7 dialogues (medium complexity); more than 7 dialogues (high complexity).

Information Density in Responses: Information density directly impacts user cognitive load and decision-making efficiency, subsequently affecting the overall user experience and satisfaction. Miller proposed that individuals can

effectively process between 5 and 9 information units (chunks) at a time in short-term memory, with an average of about 7 [6], while Nelson Cowan suggested that the working memory capacity of adults is actually closer to 4 independent information chunks [2]. Information density is divided into: low density, approximately 1–3 chunks (low complexity); medium density, approximately 4–7 chunks (medium complexity); high density, more than 7 chunks (high complexity).

Waiting Time: Following Jakob Nielsen's usability heuristic principles, there is a negative correlation between waiting time and user satisfaction [7]. Limited working memory capacity means that longer waiting times may exceed a user's ability to retain information, leading to decreased efficiency in execution [1]. Waiting time is differentiated as: no waiting required (low complexity); short and definite waiting time (medium complexity); long or indefinite waiting time (high complexity).

Error Handling Efficiency: Patience and frustration tolerance exhibited by users during failures or errors determine whether they will continue to attempt to complete a function or give up [13]. Emotional responses to operational failures can influence their overall experience. When confronted with failures or errors, users need to take action to return to a normal state [9]. Error handling efficiency is divided into: rarely makes errors (low complexity); makes errors that are easily rectifiable (medium complexity); makes frequent complex or irreparable errors (high complexity).

3.2 Assign Metric Weights

To ensure that the proposed task complexity evaluation framework accurately reflects user experience, and considering the tolerance differences of task complexity among different user groups (for instance, the difference in tolerance towards waiting time between enterprise users and individual consumers), this study utilized experimental methods to determine the weights of each core metric. This approach ensures that weight distribution is more appropriate in actual application scenarios.

This research designed a series of experimental scenarios to simulate test subjects acting as procurement personnel for company cloud resource services, undertaking three main procurement tasks: Elastic Compute Service (ECS), Relational Database Service (RDS), and Object Storage Service (OSS). The experiment recruited 10 participants with a background in cloud computing, including both male and female professionals with an average of more than five years of experience, to ensure they had sufficient understanding of the specialized terms and task objectives involved in the test.

The study adopted a random design for experimental tasks, with each main task comprising five decomposed sub-tasks that represented the various complexity levels of different metrics. The design ensured that each main task comprehensively covered all complexity levels of the metrics while maintaining a total task complexity score of 10 to control potential biases and evenly distribute experimental conditions, shown as Table 1.

Table 1. The Distribution of Metric Complexity Level Across Test Tasks

Main Task	Sub-Task1	Sub-Task2	Sub-Task3	Sub-Task4	Sub-Task5	Total task complexity score
Purchasing OSS	3pts (high)	1pt (low)	3pts (high)	1pt (low)	2pt (medium)	10pts
Purchasing RDS	1pt (low)	3pts (high)	1pt (low)	2pt (medium)	3pts (high)	10pts
Purchasing ECS	2pt (medium)	2pt (medium)	2pt (medium)	3pts (high)	1pt (low)	10pts

In the execution of the test experiments, we used the Alibaba Cloud Intelligent Robot service to develop a simulated AI conversational assistant, which was accessed and interacted with via mobile phones to fulfill predefined procurement tasks. This offered users a highly authentic CUI interaction experience. Throughout the experimental process, participants engaged in real-time interaction with the AI conversational assistant, with operation data and satisfaction ratings being collected through a screen recording system. Additionally, to ensure data collection precision and scientific rigor, the study deliberately invited two experienced user research professionals to observe the operations in their entirety and to promptly document the participants' satisfaction upon the completion of each subtask. The satisfaction data gathered during the experiment is presented in the following figure:

From the collected satisfaction data, it is observable that among the five core metrics, users are most sensitive to the increase in Complexity Level of Waiting Time, with a significant decline in user satisfaction as Waiting Time shifts to a higher complexity stage. Following that, Error Handling Efficiency and Dialogue Count also show a notable decrease in user satisfaction as the Complexity Level of these metrics increases.

After the experimental data collection was completed, a detailed analysis was conducted on the proportionate impact of increases in Complexity Level of each metric on the decline in user satisfaction. The Analytic Hierarchy Process (AHP) was applied to assess the decline ratio in satisfaction of participants as the Complexity Level of each metric increased. This assessment aimed to compare the relative importance of each metric to the CUI experience, which then determined the final weight distribution results (shown in Table 2)and established the proportionate significance of each metric.

In addition to analyzing satisfaction data, we also concurrently recorded the actual duration it took for test users to complete tasks (shown in Table 3). The data shows that as the number of operation items increases, the time users need to find the target operation items also increases, especially when the number of operation items reaches seven or more, where there is a more significant increase in task operation time. Furthermore, an increase in errors significantly extended the time it took for users to complete tasks, which further corroborates the ratio-

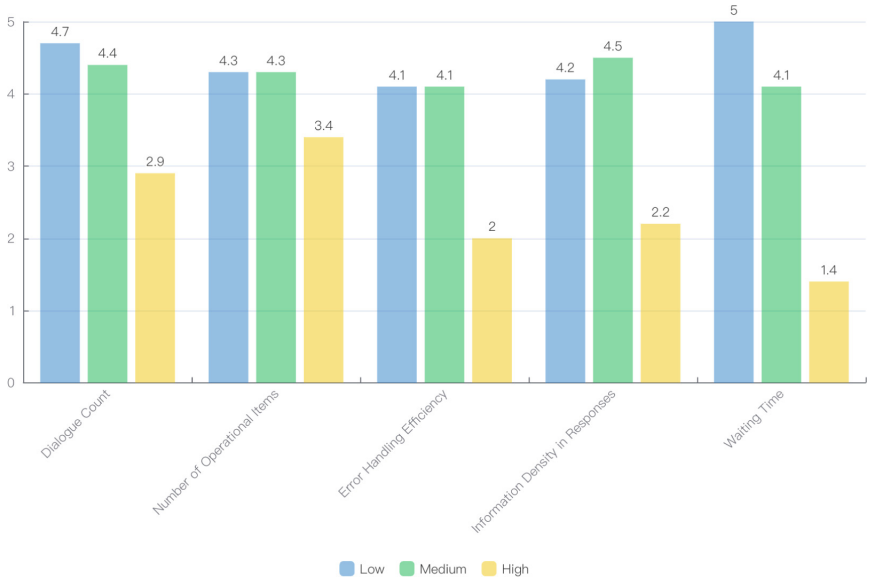


Fig. 1. Users' Satisfaction Ratings for Different Complexity Levels of the Five Metrics

Table 2. Calculated Weights Using the AHP Method

Metrics	Calculated Weights Using the AHP Method
Dialogue Count	0.19
Number of Operational Items	0.1
Error Handling Efficiency	0.25
Information Density in Responses	0.11
Waiting Time	0.35

Table 3. Execution Time of Test Users Across Various Tasks

Task Complexity Level	Dialogue Count (Dropout Rate)	Number of Operational Items (Duration)	Error Handling Efficiency (Duration)	Information Density in Responses (Duration)	Waiting Time (Dropout Rate)
Low	0%	00:13:12	00:28:57	00:39:54	0%
Medium	0%	00:18:20	01:12:19	00:36:06	10%
High	0%	00:48:42	01:17:22	01:20:02	70%

nality of our weighting definitions. These results regarding execution times are consistent with our satisfaction analysis and provide further empirical support for the accuracy of our complexity scoring and its corresponding weights.

3.3 Evaluate and Classify Complexity with Proposed Framework

Upon completing the determination, grading, and weight assignment of the metrics, this study constructed a comprehensive evaluation framework, the details of which are shown in Table 4. This framework provides a quantitative scoring mechanism for complexity in CUI tasks, allowing us to generate an aggregate complexity score. This score is designed to provide designers and decision-makers with an intuitive, comprehensible, and measurable complexity metric to identify the degree of task complexity within user experiences.

Within this framework, the first step is to define the Complexity Level Grade (G_j), which is typically represented by a set of uniformly increasing numbers presented in a linear fashion. However, if there is a need to emphasize the importance of Complexity Levels for certain metrics, non-linear values can be adjusted

Table 4. Proposed Evaluation Framework

Metrics (K_i)	Complexity Level	Complexity Level Criteria	Complexity Level Grade (G_j)	Weight(W_i)	Actual Score of Metrics (A_i)
Number of Operational Items(K_1)	Low	1–3 items	G1	W1	A1
	Medium	4–7 items	G2		
	High	more than 7 items	G3		
Dialogue Count(K_2)	Low	1–3 dialogues	G1	W2	A2
	Medium	4–7 dialogues	G2		
	High	more than 7 dialogues	G3		
Information Density in Responses (K_3)	Low	low density, approximately 1–3 chunks	G1	W3	A3
	Medium	medium density, approximately 4–7 chunks	G2		
	High	high density, more than 7 chunks	G3		
Waiting Time(K_4)	Low	no waiting required	G1	W4	A4
	Medium	short and definite waiting time	G2		
	High	long or indefinite waiting time	G3		
Error Handling Efficiency (K_5)	Low	rarely makes errors	G1	W5	A5
	Medium	makes errors that are easily rectifiable	G2		
	High	makes frequent complex or irreparable errors	G3		

as appropriate. For instance, the three levels of low, medium, and high can be defined as 0, 1, and 2 points, respectively.

Next, for each CUI task, we assess its Complexity Level for each metric, obtaining the Actual Assessment Score of Metrics (A_i), where A_i is.

By calculating the weighted sum of the Actual Assessment Scores (A_i) for each core metric along with their respective weights(W_i), we can determine the overall complexity score for the CUI application. To standardize measurements, both the Complexity Level Scores (A_i)and the weights (W_i)are normalized to a range of 1, and a ten-point system is used to display the final Complexity Score (CS). Hence, the entire calculation formula can be represented as follows:

$$CS = 10 * \sum_{i=1}^{N} A_i * W_i$$

$$A_i = \{G_1, G_2, \ldots, G_j\}$$

After obtaining the Complexity Score through quantitative analysis, we can classify the overall Complexity Level of tasks into three levels, facilitating subsequent user experience analysis and design improvements. These levels are:

Grade A (0–3 points): Represents simple tasks that provide a smooth and direct interactive experience, aligning with user expectations for the use of AI assistant applications, making them suitable for execution via CUI.

Grade B (4–7 points): Belongs to tasks of medium complexity. Although there are more layers of interaction, according to Bloom's Taxonomy of Educational Objectives, users could potentially adapt to these tasks through a reasonable amount of learning. They would only have a minor impact on the overall experience, but the experience in AI assistant applications can still be accepted by users.

Grade C (8–10 points): Complex tasks. These tasks, because of their multiple decision points and high cognitive resource consumption, may lead to significant stress and frustration for users. The experience of these tasks in AI assistant applications may not be accepted by users.

This grading system not only provides a clear framework for the design and evaluation of CUI applications but also guides developers and researchers on how to optimize task design. The goal is to ensure that interaction design is more closely aligned with users' needs and experience expectations, to enhance the usability and practical utility of the application. Through this approach, we can transform abstract complexity values into levels that have practical significance for design and user experience, providing goals and directions for the continuous improvement of the interaction experience.

We applied the aforementioned method to real business scenarios at Alibaba Cloud, conducting tests in the operational tasks of 10 mobile AI agent applications. The Complexity Level Grade we used was G (0 points, 1 point, 2 points), with weights obtained from the above experiments (0.19, 0.1, 0.25, 0.11, 0.35). We scored and classified these tasks based on the established evaluation criteria and compared our grading results with actual user satisfaction feedback on these

Table 5. Grading User Experience of Mobile AI Agent Applications Using the Evaluation Framework

Task Complexity Score	Complexity Level	Level Description	Experience Description
0–3pts	A Class	Simple Task	Meets user expectations for the use experience of AI assistant applications
4–7pts	B Class	Ordinary Task	Slightly complex operations, yet the experience within AI assistant applications is still acceptable to users
8–10pts	C Class	Complex Task	Overly complex operations, the experience in AI assistant applications may not be accepted by users

tasks. Encouragingly, the analysis results show a high correlation between our scores and user satisfaction ratings, indicating that the lower the task score, the higher the user satisfaction tends to be. Particularly, tasks graded as A (0–3 points) generally received higher satisfaction than those graded as C (8 points and above). This finding not only validates the accuracy and practicality of our method but also provides valuable insights for future optimization of the AI agent's user interface and interactions. Through the practical verification of actual business tasks, our evaluation framework has proven to be an effective tool for guiding and enhancing user.

4 Discussion and Conclusion

This research represents an evaluation framework for gauging the complexity of tasks in mobile AI agents. By selecting five crucial metrics and assigning them explicit gradations and weights, a new approach to quantitatively measuring task complexity has been established. The weights determined through experimental methods and user satisfaction adjustments allow the framework to reflect the multifaceted nature of user experiences more accurately, enhancing the accuracy and applicability of its outcomes.

In applying this framework to Alibaba Cloud's practical cases, we furthermore refine the task complexity scores into actionable tiered design strategies, which could provide invaluable insights for designers and product teams.

While our findings forge a new path in CUI complexity evaluation, the accompanying study limitations are acknowledged. The limited user sample size, although sufficient for initial insights, indicates a call for a wider data collection to solidify the universality of our results. This broader acquisition of data

would unquestionably amplify the statistical strength and application scope of our framework.

Moreover, while the weight distribution is grounded in empirical evidence, there is room for a more comprehensive capture of the diversity inherent to user experiences. Future research is thus encouraged to enrich the foundations of weight calibration. Additionally, our framework did not specifically address the impact of sensitive tasks on the complexity measurement due to time constraints; such tasks include those that may incur costs or risk data loss. We acknowledge the critical significance of these aspects across various industries, as they can considerably affect the cognitive load and emotional experience of users. Thus, future research would prioritize examining the implications of sensitive tasks and their influence on the complexity scale. Despite the identified limitations within our study, we affirm that this evaluation framework holds substantial value for enhancing the user experience associated with mobile AI agent applications. It effectively assists designers and product teams in assessing task complexity and evaluating the quality of user experience. The investigative process of this framework, coupled with the methodology of empirically calibrating parameters, not only provides a basis for experience optimization in Alibaba Cloud's mobile AI agent but also presents a universally applicable approach. It enables various enterprises and institutions to tailor the evaluation framework to their specific environments and audiences, thereby ensuring adaptability across a multitude of scenarios.

Acknowledgements. This work is part of the 2023 Beijing Industrial Design Promotion Special Fund Project_Outstanding Talent_Li Jianye, project supported by Beijing Municipal Science & Technology Commission, Administrative Commission of Zhongguancun Science Park.

References

1. Baddeley, A.: Working memory. Science **255**(5044), 556–559 (1992)
2. Cowan, N.: The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. **24**(1), 87–114 (2001)
3. Folkman, S.: Stress: appraisal and coping. In: Gellman, M.D. (ed.) Encyclopedia of Behavioral Medicine, pp. 2177–2179. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-39903-0_215
4. Hick, W.E.: On the rate of gain of information. Q. J. Exper. Psychol. **4**(1), 11–26 (1952). https://doi.org/10.1080/17470215208416600
5. Hyman, R.: Stimulus information as a determinant of reaction time. J. Exper. Psychol. **45**(3), 188–196 (1953). https://doi.org/10.1037/h0056940
6. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**(2), 81–97 (1956). https://doi.org/10.1037/h0043158
7. Nielsen, J.: Usability Engineering. Morgan Kaufmann (1994)
8. Sweller, J.: Cognitive load during problem solving: effects on learning. Cogn. Sci. **12**(2), 257–285 (1988). https://doi.org/10.1016/0364-0213(88)90023-7

9. Tice, D.M., Baumeister, R.F., Shmueli, D., Muraven, M.: Restoring the self: positive affect helps improve self-regulation following ego depletion. J. Exper. Soc. Psychol. **43**(3), 379–384 (2007). https://doi.org/10.1016/j.jesp.2006.05.007
10. Vohs, K.D., Baumeister, R.F., Schmeichel, B.J., Twenge, J.M., Nelson, N.M., Tice, D.M.: Making choices impairs subsequent self-control: a limited-resource account of decision making, self-regulation, and active initiative. In: Baumeister, Roy F.. (ed.) Self-Regulation and Self-Control: Selected works of Roy F. Baumeister, pp. 45–77. Routledge (2018). https://doi.org/10.4324/9781315175775-2
11. Wolfe, J.M.: Guided search 2.0 a revised model of visual search. Psych. Bull. Rev. **1**, 202–238 (1994)
12. Wood, R.E.: Task complexity: definition of the construct. Organ. Behav. Hum. Decis. Process. **37**(1), 60–82 (1986). https://doi.org/10.1016/0749-5978(86)90044-0
13. Zeelenberg, M., Pieters, R.: Feeling is for doing: a pragmatic approach to the study of emotions in economic behavior. In: Social Psychology and Economics. Psychology Press, pp. 117–137 (2013)

Author Index

A
An, Qi 27, 210
Anbalagan, Sanjay Krishna 223
Aristi, Savira 232

B
Biłyk, Anna 332
Braun, Marie 243
Buele, Jorge 406
Bullard, Enya 251

C
Chen, Jiaxin 117
Chen, Xiaoli 3
Chen, Yazhou 15
Chen, Yu 285
Cheng, Theodore H. 256
Cheng, Xin 48
Chincholikar, Kartik 152
Chacon Quintero, Juan Carlos 232

D
de Los Santos, Sabrina 436
Dong, Fangzhou 144
Dübeler, Dennis 263
Dwivedi, Shagun 152

E
Elsayed, Yomna 269
Essing, Britta 368

F
Feger, Marc 243
Fu, Boxi 436
Fulbright, Ron 280

G
Goh, Zhang Hao 377
Gopalan, Kaushik 152
Greenwald, Eric 436
Guo, Chang 58

H
Ha, Michelle 285
Hammerman, Jim 436
Hering, Maike 294
Hernández-Alvarez, Myriam 304
Hernández-Ramírez, Rodrigo 48
Hespanhol, Luke 48
Huang, Xinyu 27
Hunger, Robert 312
Hurt, Timothy 436

I
Ichikawa, Jun 323
Ichinose, Genki 323
Ikeda, Kazushi 412
Iwaihara, Yutaka 323

J
Jiang, Lu 37
Jin, Guo 106

K
Kanamarlapudi, Vijay Kumar 223
Khan, Nibraas 251
Kobayashi, Nao 412
Kommalapati, Anughna 223
Krakowski, Ari 436

L
Łabędzki, Rafał 332

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

C. Stephanidis et al. (Eds.): HCII 2025, CCIS 2529, pp. 455–457, 2025.
https://doi.org/10.1007/978-3-031-94171-9

Lai, Zhuosi 48
Lee, Chei Sian 349, 359, 377
Li, Anglu 58
Li, Chenyu 444
Li, Jianye 444
Li, Li 117
Li, Xinnan 76
Li, Yichen 128
Li, YouLan 89
Liang, Na 89, 172
Liao, Jiaqi 349
Liao, Keyan 3
Liao, Yanting 106
Lim, Kok Khiang 359
Lindberg, Susanne 386
Ling, Jie 117
Liu, Jia 27
Lu, Zhiyuan 210
Lund, Jesper 386

M
Martinez Baca, Alejandro 232
Martinez Nimi, Hisa 232
Masur, Omkar 436
Merchant, Chirag 436
Mikołajczyk, Katarzyna 332
Miyamoto, Akihiro 412
Mohan, Umesh 223
Morad, Milad 368
Müller, Alissa 263

N
Nakajima, Kae 412
Naruse, Yasushi 412
Neef, Nicolas 294
Nei, Xinrui 223
Nguyen, Chi Tran Mai 377
Norris, Simon 386

O
Ono, Kenta 232
Otto, Siegmar 294
Ou-yang, Li 117

P
Patiño-Fernández, Rocio 406
Pérez-Vega, Doris 406

Q
Qiu, Keyi 397

R
Resmini, Andrea 386
Richter, Birte 263

S
Sanada, Motoyuki 412
Sarkar, Nilanjan 251
Schütze, Christian 263
Song, Chenjun 37
Song, Yutong 128
Subía-Arellano, Andrés 406
Sugisaki, Eriko 412
Sun, Shenghang 27
Sun, Yijia 144
Suwa, Hirohiko 424

T
Tahara, Shunichi 412
Tandoc, Edson C. 377
Theis, Sabine 312
Torres-Hernández, Edgar 304
Trojanowska, Monika 332

U
Ueda, Yuji 424

V
Valaboju, Bharath 152
Vidwans, Vinod 152

W
Wang, Bin 89, 172
Wang, Honglei 27
Wang, Jiamin 37
Wang, Jian Feng 172
Wang, Junyi 37
Wang, Ning 436
Wang, Ran 117
Wang, Ruoxuan 3

Wang, Ziyang 161
Wen, Jiazhi 444
Wrede, Britta 263
Wu, Weilong 185

X
Xie, Junci 117
Xu, Dawei 195

Y
Yamada, Masatoshi 323
Yang, FangChao 172
Yang, Yani 210
Yang, Yuwei 444
Yasuda, Keiji 412

Z
Zabel, Sarah 294
Zeng, Yue 128
Zhan, Jing 3
Zhang, Huafang 117
Zhang, Tianrun 397
Zhao, Xiaodan 223
Zhao, Yi 48
Zhao, Yuanfang 76, 161
Zhong, Cuiyan 117
Zhou, Zhanyan 37
Zou, Jingjing 185