Challenges and new directions for the certification of AI and advanced automation in civil aviation

Paola Lanzi, Elisa Spiller
Deep Blue
Rome, Italy

Mohsan Jameel, Lothar Christoffels
Institute of Flight Guidance
DLR
Braunschweig, Germany

Gabriella Gigante, Riccardo Esposito
AI and Complex Systems Unit, Cross Cutting Research Direction – CIRA
Capua, Italy

Giuseppe Contissa, Marco Sanchi, Federico Galli
Department of Law
EUI
Florence, Italy

Mariken Everdij, Sybert Stroeve
Safety & Human Performance
NLR
Amsterdam, The Netherlands

***Abstract*—This paper examines the certification challenges and opportunities posed by advanced automation and AI-based systems in Air Traffic Management (ATM). The findings presented are the outcome of research carried out in the first part of the SESAR Joint Undertaking's HUCAN project, which aims to develop a holistic, unified certification framework for highly automated systems, addressing both technical reliability and human factors, to ensure their safe and efficient integration into ATM. Based on key issues from ongoing technical and regulatory discussions, the research evaluates current innovative certification approaches and explores how levels of automation can enable a more refined development and certification process. In light of the results obtained, the paper proposes new directions that could advance a holistic approach to certification, considering the guidance provided by EASA.**

Keywords—advanced automation; artificial intelligence; certification approaches; civil aviation; human-computer interaction; human-AI teaming

I. Introduction

Over the past decade, the technology landscape has undergone a remarkable transformation, paving the way for the expanded role of automation. One of the most significant developments has been the evolution of these systems' capabilities to analyze complex situations, learn from them, and make intelligent decisions autonomously. In some cases, this progression has exceeded expectations, as high automation systems have demonstrated proficiency in cognitive tasks [1] previously believed to be exclusive to human capabilities.

The field of Air Traffic Management (ATM) has seen rapid changes and growth, with higher levels of automation integrated in ATM systems to support the human operators in information acquisition and exchange, information analysis, action selection and action implementation for all tasks/functions. By leveraging highly automated systems, the ATM industry can potentially transform how it operates with improvements in safety, efficiency, and reliability [2]. Automated systems can analyze large amounts of data from various sources, including radar, weather sensors, and flight plans, to provide controllers with real-time insights and decision support. With the use of machine learning (ML) algorithms, these systems can adapt to changing conditions and optimize air traffic flow dynamically, leading to smoother operations and fewer delays. Additionally, this technology can assist air traffic controllers and pilots by reducing their workload and alleviating stress associated with their responsibilities. Also, increasing the levels of automation in ATM systems can augment human capabilities rather than replace them entirely, which is clearly expressed in EASA Artificial Intelligence (AI) AI Roadmap [3] as a human-centric approach.

Despite the clear benefits and progress made in automation technology, there remain significant challenges surrounding the integration of higher levels of automation and AI in ATM operations. These challenges encompass concerns related to human-machine interactions and the organizational impact of automation [4] [5]. Also, paramount to the success of these advanced systems is the assurance of security and resilience against cyber threats [6]. Striking a delicate balance between automation and human involvement, especially in decision-making processes, poses a multifaceted challenge [7].

The SRIA for Digital European Sky [2] also identifies key issues and challenges, particularly the need to focus on developing new methodologies to validate and certify advanced automation that ensure transparency, legal compliance, robustness, and stability under all conditions. This should be done with consideration of a future ATM environment that relies on multiple AI-based systems of systems, with a focus on human-centered design.

The SESAR JU project "HUCAN - Holistic Unified Certification Approach for Novel systems based on higher levels of automation" aims to develop a novel approach for the certification and approval of innovative ATM-related airborne and ground systems based on advanced human-centric automation, including those based on AI-powered solutions. The focus on human-centered automation is key in this framework, as it implies the need for a holistic approach to certification that, together with technical reliability, takes into consideration also other aspects of the system, such as its impact on human performances and operations (including human-system teaming and training and skills requirements), and at societal, value-based and ethical level. In addition, the proposed certification approach intends to be unified, meaning able to take into account and fit the wide range of deterministic and non-deterministic algorithms available and the automation of tasks and procedures according to different levels of automation and autonomy. The novel approach will support both the approval/certification process and the design phase of highly automated technologies, proposing to this end the development of two interconnected and tangible products: a new holistic and unified certification method for systems based on advanced automation, and a set of suitable guidelines and associated toolkit for streamlining the development of highly automated and AI-based technologies.

In this paper we present the output of the research carried out in the first part of the project, exploring in particular the current state of the art that sets the stage for the design of the two innovative products above. To this end, we first analyze the needs and challenges that exist in the ATM sector for the implementation of advanced automation, with the aim of gaining a contextual understanding of the requirements for standardizing regulatory frameworks and adapting certification methods to ensure the safe and responsible use of advanced automation technologies (Section II). We then present an overview of the emerging and innovative approaches to the approval and certification process of automated and AI-based technology for the domain of civil aviation, opportunely reviewed according to a set of evaluation criteria that measure their applicability for advanced automation (Section III). Based on this, we conclude the paper by outlining a set of new directions in the certification of AI and advanced automation for civil aviation that will allow to address the challenges and limitations associated to the approaches analyzed and that will be used to drive the definition of the HUCAN innovative, holistic and unified approach (Section IV).

II. BACKGROUND AND CHALLENGES

A. *Challenges of Innovation in automation for civil aviation*

Nowadays, the integration of advanced automation and AI seems to be a revolution, being recognized as a fundamental force to bring safer, resilient, ecological and efficient solutions in the domain driving the industries' evolution including transport management solutions [8]. In the aviation domain, advanced automation has improved the management of different aspects of air traffic operations, integrating task automation, leading to a substantial enhancement of safety, efficiency and accessibility of air transport. It can positively impact on the safety of both manned and unmanned aircraft, including Unmanned Aerial Vehicles (UAVs) and drones. In the area of Urban Air Mobility (UAM), advanced automation and AI applications are making strides in improving route planning techniques. Projects such as LABYRINTH [9], SAFEDRONE [10], MONIFLY [11], COMP4DRONES [12], TINDAIR [13], and AURORA [14] are crucial for optimizing route planning, increasing the efficiency and safety of urban air transport. These technologies also contribute to more precise positioning of drones and UAV within the U-space (GAUSS [15]), which is crucial for maintaining order and safety in increasingly crowded airspaces. Moreover, enhancements in sensor performance and connectivity protocols are being achieved through initiatives like SAFEDRONE [10], SAFIR-MED [16], and ASSURED-UAM [17].

Significant progress is also expected in ATM, especially through the integration of AI and digital tools designed to assist Air Traffic Control (ATC) decision-making. AI-driven applications were previously explored in projects MAHALO [18], PJ16 CWP HMI [19], FARO [20] and a stream of on-going projects such as JARVIS [21], DARWIN [22], and LOKI [23] provide major benefits by easing decision-making processes, reducing the workload of controllers and allowing operators to concentrate on critical tasks by automating routine procedural tasks. Predictive modelling techniques are being actively studied to predict safety risks and collect operational data from flights. These efforts highlight the importance of storing and processing large volumes of operational data to identify patterns and train predictive algorithms, thus leading to greater safety and efficiency in ATM (SAFECLOUDS) [24].

AI is being used to support the Dynamic Airspace Reconfiguration (DAR) process, to provide optimal trajectory, assuring safety, minimizing the human workload and the environmental impact.

In-flight safety research focuses on the use of advanced sensor technologies and software processing techniques to improve overall safety. Projects like PJ11 CAPITO [25] and ODESSA [26] envision the establishment of future safety requirements. In particular, different projects are obtaining interesting results in obstacle detection, avoidance, and navigation, especially under challenging conditions such as low visibility or adverse weather (SENSORIANCE [27], WINFC [28], VISION [29]). Also, there is a coordinated effort to monitor the cognitive state of pilots and assess the impact of highly automated systems on ATCOs performance (STRESS [30], REPS [31]).

In emergency scenarios, significant research and development efforts focus on enhancing pilot decision-making during emergencies (SAFENCY [32]) and managing scenarios involving onboard pilot incapacitation (SAFELAND [33]). These research initiatives aim to improve rescue capabilities in general aviation emergencies, contributing to a holistic approach to safety management in aviation emergency scenarios (GRIMASSE [34]). In parallel to such evolution, Human-Machine Interface (HMI) design is becoming crucial to facilitate collaboration with AI assisted decision-making environments (PJ16 CWP [19], HAIKU [35]).

The analysis of the research already carried out reveals that despite the numerous opportunities AI and advanced automation offer, several challenges are rising, in order to ensure safe and effective integration into aviation systems.

1) *Safety, Reliability, and Robustness of Systems*

- It is crucial to ensure that the new systems do not perform any unintended function.
- Safety can be jeopardized by the level of generalization achieved by the AI-algorithm. The way it extrapolates its knowledge from new data which might differ from the original training data, represents today a key research topic.
- AI learning capability and the required robustness leverage also on data assurance. Services to process aviation data and related infrastructures could represent further opportunities in the next future.
- AI requires further refinement to handle uncertainty when making predictions.

2) Human-Machine Interaction and Teaming

- To assure a smooth collaboration between human and its advanced assistant, Human-Machine Teaming is becoming a key research stream. This includes dynamic task division, intuitive interfaces, and a multi-disciplinary approach for social science for AI experts.
- The design of the HMI must facilitate seamless collaboration and cooperation, ensuring clear and intuitive interaction.
- The design of the Human-Machine Teaming requires intuitive and clear user interfaces.
- In crisis management, there is a need for shared situational awareness between the human and its AI assistants to ensure effective decision-making.
- The relationship between humans and machines must be carefully mapped to enable complex and effective cooperation.

3) Transparency, Explainability, Trust and Ethics

- Transparency and explainability of AI's decision-making processes are crucial for building trust since the reasoning process and the solutions presented to the human operators must be clear so that they can trust the AI outcome and can take well-founded decisions.
- Ethical challenges must be addressed in order to make use of controversial but powerful techniques like profiling, which could optimize human-assistant tuning but also present new challenges and opportunities.

4) Regulation, Standards, and Cross-Domain Harmonisation

- The power of continuous learning of AI triggers the definition of new safety precautions, potentially leading to new regulations and roles.
- There is a need for a harmonized AI and advanced automation related taxonomy across all transport domains, to maximize research potential through knowledge exchange.
- The scientific community and the relevant stakeholders should maximize the potential of the research by means of cross fertilization.

B. Challenges in Current Certification Methods

Certification is defined as any form of recognition, based on an appropriate assessment, that a product, part or appliance, organization or person complies with the applicable regulatory requirements, through the issuance of a certificate attesting such compliance (EU Reg 2018/1139). The certification is aimed to be the proof that something has been thoroughly analyzed and tested and can be safely used.

In Europe (focus of the present research), the requirements and certification standards are set by EASA (European Union Aviation Safety Agency), and include rules on certification of the operator, the licensing of pilots, certification standards for an aircraft, and required processes for developing software. The demonstration of compliance may take years and a significant number of resources. The methods and tools used take into account international, national and European technical standards from authoritative organizations and bodies, formed through consensus and then rationalized, collected and published. As an example, in the development phases of aircraft, the demonstration of compliance includes applying safety assessment methods such as SAE ARP4761 [36] (equivalent to EUROCAE ED-135 [37]) for hardware, and RTCA standard DO-178C [38] (equivalent to EUROCAE ED-12C) for software. This is followed by ground testing (such as tests on the structure to withstand bird strikes, fatigue tests and tests in simulators) and flight testing by licensed test pilots in a variety of circumstances.

ARP4761 has limitations when being applied to advanced automation and AI-powered technology, since these require the ability to study dynamic scenarios, the complex interaction with the environment, feedback loops and non-linear interactions, etc., rather than component failures. The method focuses on functional failures rather than on other hazards. The modelling techniques used as part of the ARP4761 process, such as fault trees and failure mode and effect analysis, are not suitable for the analysis of human performance or procedural aspects since they are not sufficiently able to deal with dependencies and dynamics. In the development of AI-supported systems and operations, it is expected that there will be a shift in the level of authority towards increasingly autonomous systems with decreasing in-the-loop roles of human operators. This implies that a part of the intelligent contributions of human operators are taken over by contributions of AI-based systems and that the roles and responsibilities of human operators change. Clearly, such a shift has considerable legal implications for the allocation of responsibility and liability for product and service developers versus the organizations and people providing services and using the products.

Safety is not an intrinsic property of an AI-based system. The safety impact of a particular AI component in a system depends on the dynamic interactions with other systems, humans working with operational procedures, and contextual conditions. The traditional safety assessment approaches have their origins in assurance schemes for physical components, which may fail/break and for which statistical quality control approaches can be applied. These approaches are known to have limitations for assessing and controlling the safety impact of software, AI-based systems, and human factors (HF). While these limitations also apply to systems and operations without AI components, their implications may be aggravated due to the new aspects brought by ML and due to the shift in contributions of humans and AI-based systems in the operations.

III. Innovative Certification Methods

This section presents innovative certification approaches from the domain of civil aviation, AI and advanced automation. Next, the section examines the validity of these approaches by introducing a series of evaluation criteria centered around a broad and renewed understanding of the goal, scope and means of a certification approach, including a discussion regarding the use of automation levels in the domain. Based on such criteria, it highlights challenges and limitations emerging from the assessment, and shares ideas, future research and new directions for the certification domain.

A. Innovative Approaches in AI and advanced automation

EASA maintains its pivotal role by spearheading efforts to integrate AI into civil aviation through its comprehensive AI roadmap and guidance for ML applications [3] [39]. This roadmap prioritizes the development of a human-centric and trustworthy AI, confirming EASA as a leading certification oversight authority in the domain. By supporting EU leadership in AI, aligning with broader EU strategies, such as the provisions of the AI Act (AIA, Reg, (EU) 2024/1689), and fostering a robust research agenda, the roadmap aims to ensure that AI systems deployed in aviation are safe, explainable, and aligned with human factors. Additionally, it offers organizational recommendations for agencies seeking to implement AI, with a strong emphasis on Explainable AI (XAI) for operational purposes, showing a broader view regarding the role of certification in the field.

Complementing these efforts, the recently enacted and previously cited AIA addresses the challenges posed by AI development and deployment as a general and encompassing framework, introducing regulation relevant for a variety of domains, including that of civil aviation and advanced automation (Reg. (EU) 2024/1689). The new legislative framework characterizes the AIA, which follows a risk-based structure, classifying AI systems into risk levels or classes based on their potential risks to safety, health and fundamental rights, and defining tailored legal regimes for each category. By promoting trustworthy and human-centric AI, the AIA parallels efforts by EASA and ensures that AI systems are tools that uphold safety, health, and fundamental rights, by introducing a robust monitoring and enforcement framework to mitigate their risks, fostering their application in the aviation industry as a safety critical one.

Ethical considerations appear to be considered as integral and critical to the certification and deployment of AI, referencing the European Commission Ethical Guidelines for Trustworthy AI [40] [41]. The guidelines tie the notion of trustworthiness to that of lawfulness, ethics, and robustness, ensuring that AI systems are deployed in a fair and explainable way, respecting the need for human autonomy and parallelling already tackled approaches.

In addition to these regulatory efforts, standardization initiatives contribute to shaping the certification landscape. The EUROCAE Working Group 114, in collaboration with SAE G-34, is developing a standard for the certification of aeronautical safety-related products that implement ML [42]. This effort includes a comprehensive gap analysis that identifies limitations in existing standards and focuses on offline learning applications, where ML models are trained and deployed in fixed systems. Their work on standardization appears focused on design assurance levels (DALs) for technical systems and of technical value.

Furthermore, ISO/IEC standards appear to play a pivotal role in the certification of AI systems in aviation, with a focus on ensuring product quality, robustness, reliability, and safety. These standards, such as ISO/IEC TR 5469:2024 for functional safety [43], ISO/IEC CD TS 6254 for AI explainability [44], and ISO/IEC 25000 for overall system quality [45], show a willingness to address the unique challenges posed by AI while leveraging existing best practices in automation, in an effort to deploy safe and trustworthy AI systems, aligning with broader goals of transparency, reliability, and maintainability.

Finally, the IEEE's CertifAIEd program further enriches the picture by providing a comprehensive framework for AI ethics certification, covering transparency, accountability, algorithmic bias, and privacy [46]. The program includes so-called ontological specifications, with the goal of offering insights into these criteria and ensuring compatibility with emerging regulations, considering harmonization among certification approaches and standards. Also, the CertifAIEd ecosystem values the inclusion of trainers, assessors, and certifiers who play crucial roles in disseminating knowledge, assessing AI systems, and certifying compliance with ethical standards, paying attention to the organizational and enforcement dimension of certification.

Research shows that the certification domain of AI, advanced automation and civil aviation is rapidly evolving, taking on a broader view of certification as a process going beyond technical safety, emphasizing the integration of human-centric AI and trustworthiness frameworks, while additionally focusing on lawfulness, ethics, and robustness. The role of industry standards still focuses on ensuring safety, and reliability, though the need for explainability and concerns of harmonization are additionally underscored.

B. Innovative Certification Approaches Assessment Criteria

In an effort to define evaluation criteria, information is drawn from authoritative sources and frameworks from the aviation industry, such as the SESAR Performance Framework, the S3JU Multiannual Work Programme 2022-2031 [47], and the European ATM Master Plan 2020 [48], as well as the already cited Roadmap on AI from EASA.

Moreover, widening the scope of analysis, additional sources are gathered from outside the boundaries of the civil aviation field, expanding it to include criteria and features capable of thoroughly assessing the suitability of certification approaches dealing with the implementation of AI and advanced automation in the domain. In this regard, the Digital Decade Policy Programme 2030 [49] and the AI for Europe Strategy [50] are included, as well as input from the preliminary objectives and guidelines of the S3JU, as outlined in the pre-read material of the European ATM MP Stakeholder Consultation Workshop of 8 April 2024 [51].

The goal is to support the use of criteria capable of scrutinizing the effectiveness of state-of-the-art certification approaches for application to AI and advanced automation, as well as the combination of the two, in civil aviation. Therefore, a holistic approach is adopted, scrutinizing certification frameworks under a variety of different, yet interconnected viewpoints, to achieve a thorough and comprehensive analysis. Moreover, given the intrinsically intertwined nature of some of these critical assessment aspects, criteria may present elements that overlap, such as that of explainability, relevant in terms of uncertainty as much as of human factors, or technical complexity. The analysis takes this into account and presents assessment results as a discussion relating to the application of all the criteria as a whole. The list of assessment criteria produced by this research includes the following:

Uncertainty. Evaluate how the certification approach addresses uncertainties in technology, data, operational scenarios, and unforeseen events, especially for high automation levels. It also assesses contingency planning for major failures and security breaches.

Safety. Assess the certification approach's support for comprehensive risk management, including safety and security interfaces. It examines the identification of risk indicators, risk assessment detail, and compliance methods for safe operation.

Accountability. Measure the effectiveness of the certification approach in defining responsibilities across the aviation value chain, balancing stakeholder discretion with consistent safety standards, and ensuring ongoing compliance with established frameworks.

Environmental Protection. Evaluate how well the certification approach supports the reduction of air travel's environmental impact, while also fostering the adoption of international environmental standards.

Public Oversight. Assess the extent of democratic control and public participation in the certification process, ensuring transparency, independence, and public interest are maintained, even when certification duties are delegated to private entities.

Efficiency. Evaluate the balance between innovation and regulatory rigor in the certification process, ensuring safety without hindering technological advancement, and assessing the time required to complete certification.

Technical Complexity. Measure the accessibility and clarity of the certification approach, including the explainability of its application and the complexity of the tools required, ensuring it is manageable for experts without excessive resource demands.

Human Factors. Evaluate the consideration of human interaction with automation, focusing on human-AI teaming, training programs, and practices that promote safety culture, situational awareness, and responsible system reliance.

Data Governance. Assess the ability of the certification approach to establish robust data governance, ensuring data accuracy, safety, and accessibility. It includes protocols for data access, storage and use, adhering to relevant regulations, and for the prevention of biases.

C. Assessment Results: Challenges and Limitations

The application of the above cited criteria to innovative certification approaches highlights some limitations in the current frameworks for AI and advanced automation in civil aviation. This section presents the results of the assessment carried out so far, arguing for the need to further develop the human element of certification, to consider essential socio-technical factors such as accountability, oversight and human control, as well as to present the impact that asymmetry of information has on the certification process. Finally, it builds on the identified gaps to discuss implications for future certification strategies, suggesting new directions and approaches.

The assessment results for each criterion are:

Uncertainty. Most approaches recognize the importance of managing uncertainty in advanced automation, but there is scope to further explore and address these issues in relation to human interaction with technology.

Safety. The approaches analyzed prioritize the impact of technical aspects of advanced automation on safety and there is room for further research and investigation, also taking into account the impact of organizational aspects and HF on safety.

Accountability. Accountability is considered, but it should be given greater emphasis by integrating it directly into architectures and the certification process to ensure clear alignment with technical standards.

Environmental Protection. Currently, few certification frameworks address environmental protection directly, relying more on external initiatives and processes. However, this topic should become a more crucial focus, similar to accountability.

Public Oversight. The role of public oversight and stakeholder engagement in the certification process is considered, yet only partially implemented. This suggests the need for more oversight of certification agents and increased societal participation.

Efficiency. Efficiency is given limited consideration due to a trade-off with safety, requiring a careful balance. Moving forward, efficiency in certification should be prioritized given their significant impact on overall performance.

Technical Complexity. The high technical complexity of advanced automation presents opportunities for broad stakeholder engagement to manage its challenges. This should be better considered and implemented in the future.

Human Factors. Human and ethical factors such as explainability, agency, and trust are considered, promoting human-centred AI approaches. Moving forward, it is essential to provide clear criteria for enforcing binding standards that take these aspects into account.

Data Governance. The topic needs further exploration, especially given its importance in AI and automation applications. There is an opportunity to enhance standards for data management and cybersecurity, and to implement these standards directly within processes.

To summarize, one of the most significant findings from the assessment is underscoring the presence of a "human gap", that is, insufficient integration of HF in the certification process. In fact, current approaches appear to treat HF, such as trust, accountability, and human-technology interaction as abstract principles or guidelines, rather than implementing them as concrete elements systematically incorporated into the certification process. This omission creates a critical limitation, as the human element plays a pivotal role in ensuring the safe and effective operation of AI and automation systems in the aviation domain.

In this regard it can be noted how, while safety seemingly occupies a central role in all innovative certification approaches, its focus nevertheless predominantly lies on technical aspects and product safety, often not, or not properly, addressing the broader socio-technical context within which automation operates. The interaction between humans and AI systems, with particular regard as to how decisions are taken and by which actors, as well as the potential for automation bias, seems to be addressed inadequately. What follows is the potential cause of scenarios where human operators are either overly reliant on automated systems or insufficiently prepared to intervene when necessary, compromising overall safety and confirming the pivotal importance of the human element.

Another critical issue identified is in the asymmetry of information between the various stakeholders involved in the certification process. When analyzing certification approaches, the assessment shows that the actors involved in the certification process, which includes, amongst others, regulators, developers, and operators, do not seem to have equal access to relevant information. Technology providers, particularly those developing AI systems, emerge as actors possessing a higher degree of practical information and detailed knowledge about the capabilities, limitations, and potential risks of AI models, particularly in the context of advanced automation, than regulators or operators. Following all of the above, this inadequate balancing of information may lead to oversight complications, including critical uncertainties and lackluster accountability frameworks surrounding the certification process. Moreover, asymmetry of information additionally affects public oversight, which is inconsistently addressed across certification approaches.

Although effective stakeholder participation and oversight by competent authorities are crucial for ensuring transparency and trust in the certification process, the assessed approaches often fail to provide a structured mechanism for involving a broad range of stakeholders in the certification process. There is an additional tendency to focus on the role of technical standards for certification, disregarding the role and concert played by relevant actors and the procedures involved in the certification process.

All of the above additionally ties into further discussion surrounding the scope of innovative certification approaches. In fact, the assessment suggests that the focus of certification appears to be narrow and of limited scope, particularly in their treatment of safety, human elements, and socio-technical processes, as well as environmental protection. While safety is rightly prioritized, it is in fact generally considered in a static and technical sense, without sufficient attention to the dynamic and evolving nature of AI systems, including aspects of human-computer interaction, agency and control. This narrow focus limits the effectiveness of certification in ensuring the long-term safety and reliability of AI systems, especially in a rapidly changing technological landscape.

Moreover, environmental protection is often relegated to external standards or legislative frameworks, rather than being integrated directly into the certification process, in a similar fashion to accountability and public oversight. Given the growing importance of sustainability in all sectors, including aviation, this represents a significant gap, which moves us to argue for certification frameworks to incorporate environmental considerations, information and provisions as a core component of their process, rather than relegating their integration to external third sources of regulation.

Above all considered, through what is described as a holistic approach to the issue of assessing innovative certification approaches, the presence of a series of intertwined and pervasive gaps, challenges and limitations is highlighted, with numerous emerging patterns. The next section builds on the results discussed thus far to present additional findings regarding the interplay between automation, AI and the human element, as well as their impact on certification.

D. Automation Levels with AI as a key enabler

The gaps identified in the review of the state of the art on innovative certification approaches are consistent with the directions that are driving ongoing institutional and industry discussions on high levels of automation.

Technologies enabling higher levels of automation in practice involve a significant redistribution of operational authority between operators and systems, with potentially serious implications for the safety and accountability of frontline operators and their organizations. The feedback collected by EASA and SESAR highlights the urgent need to rethink and update the certification process for advanced automation in the light of the profound impact of these new operational paradigms on human-machine interaction, taking into account the interdependencies of responsibilities along the value chain. Current certification frameworks prioritize safety and technical standards but are inadequate for addressing the complex socio-technical challenges brought by these technologies.

In response to these challenges, the role of levels of automation within different approaches becomes particularly important and serves as a critical driver for redefining certification requirements and processes. Based on the principle of proportionality, the assumption is that certification approaches for advanced automation and AI should be designed and adapted according to the level and scope of risks that systems may pose, rather than on techniques used to develop the AI-based systems. While these solutions have distinctive and disruptive features that impact both operational safety and cybersecurity, they should be evaluated based on the specific functions they perform within a given process, considering the scope, frequency, and criticality of these tasks.

These considerations align closely with the risk-based approach guiding the regulatory discussion on AI and advanced automation in aviation, leading EASA and SESAR to provide valuable guidance for addressing the challenges in AI and advanced automation research, development, and certification through a broader and collaborative approach.

The EASA Roadmap has made AI classification a cornerstone of its AI trustworthiness building blocks [39]. This taxonomy distinguishes systems based on the contribution they make and the authority they assume when interacting with human operators and the environment. Applications can be clustered into three levels, distinguishing assistance to human (Level 1), human-AI teaming (Level 2) and advanced automation (Level 3).

Complementarily, in the preparatory material for the review of the European ATM Master Plan, SESAR proposed integrating the AI levels suggested by EASA with its traditional LOAT [51]. This approach is generally intended for advanced automation solutions and spans a six-level system, ranging from low automation to full autonomy without human supervision. The aim of this different approach is to converge attention on the level of automation achieved by the different applications considering the impact of the technological innovations on the main human cognitive and operational functions.

HUCAN provided the opportunity for a first experimental application of these taxonomies on four use cases concerning innovative operational concepts and systems for dynamic use of airspace and capacity on demand.

The analysis highlighted the level of automation and the human factors impact of the reference systems, providing a preliminary analysis of the gaps and challenges in the evaluation of the level of automation and the liability assessment in practice. In particular, the relevant issues that have emerged relate to both potential substantive and procedural certification issues.

The results reveal that, as first, the introduction of highly automated solutions, including AI, has significant impacts on human roles, tasks, and tools, requiring adjustments in responsibilities and interactions with the system. Second, there are ambiguities in the level of automation classification for automated systems, particularly concerning decision support and detection/prediction functions, which indicate gaps in the current framework. Finally, automation, especially AI, raises liability concerns for developers, operators, and regulators. These risks are particularly relevant in concepts involving multiple operational modes, where dynamic transitions between full automation and human-machine collaboration occur, urging the need for clear protocols and timely human intervention to prevent safety issues.

Moving on to the considerations that emerged regarding the certification process, highly automated technologies should be certified based on their specific operational use, which may vary from location to location, requiring tailor-made certification for each application rather than a one-size-fits-all approach. All automated technologies, including AI, share common certification challenges, highlighting the need for a unified certification process that addresses issues such as explainability. In addition, early assessments at lower maturity levels are essential to identify potential HF and liability issues, and to develop appropriate mitigations, such as role adjustments or changes in the level of automation.

IV. Conclusions and Future Works

The research conducted so far in HUCAN indicates that, considering concerns about high automation and AI in aviation, the current certification system requires revision. Beyond the technical focus and high degree of attention to safety, there are identified areas where enhancements are needed to more effectively address the broader socio-technical challenges associated with these technologies.

As noted, these challenges include concerns related to human-machine interaction and the organizational impact of automation. However, the need for new approaches to certification goes beyond the deployment of solutions and fundamentally affects the entire design process. As shown, early analysis of concepts from a holistic, human-centered perspective helps to assess the benefits and risks associated with different levels of automation and to explore alternative design options, allowing for adjustments in responsibilities and interactions with the system. In addition, it is important to recognize that not all the identified critical areas can be addressed immediately; they should be taken on board progressively as the concept and technologies reach an appropriate level of maturity.

Considering the substantial and procedural certification issues identified by the analysis of the state-of-the-art, HUCAN identifies three new directions to pursue to achieve these goals.

Firstly, the level of automation is a key factor in driving holistic certification processes. Therefore, to improve the approach to certifying advanced automation systems, it is critical to adopt a phased, holistic certification strategy that leverages the insights gained from the proactive and diligent application of these categories. This strategy allows for more effective risk assessment in practice. Integrating these taxonomies into new certification frameworks will help to develop a process that is both proportionate and comprehensive, ensuring that operators are held to an appropriate level of accountability.

More attention needs to be paid to HF in new certification processes, assessing how well human operators can understand, trust and control AI systems, considering the ethical and societal expectations and risks associated with different levels of automation. It is essential to establish enforceable standards rather than relying on guidelines alone. In addition, especially in systems with varying levels of automation, clear protocols and timely human intervention are required to effectively prevent safety issues and promote a proactive, mitigative approach to liability concerns related to automation and AI.

Finally, it is important to consider the operative and social expectations regarding the certification process, verifying the opportunity to extend the scope and include new dimensions and methodologies. Critical socio-technical aspects such as accountability, public oversight and environmental protection are of crucial importance here. It is essential to draw on expertise in different fields and at different stages of the development process to refine appropriate certification approaches. This should include establishing clear decision-making roles, enforcing sustainability standards, and using public oversight and stakeholder input to enhance transparency and trust in a collaborative and transparent manner.

An approach that includes these aspects not only addresses the specific needs of new solutions by adapting the regulatory and certification regime to their actual characteristics and risk levels, but also embodies a co-creative process. Close cooperation with EASA and SESAR will be essential to continuously refine these levels, both in theory and in practice, and to achieve the desired objectives.

REFERENCES

[1] OpenAI, ChatGPT (GPT-4) Large Language Model, 2024.

[2] SESAR JU, “Digital European sky – Strategic research and innovation agenda”, Publications Office of the European Union, 2020.

[3] EASA, “Artificial Intelligence Roadmap 2.0: Human-centric approach to AI in aviation”, May 2023.

[4] W. M. Lim, "The workforce revolution: Reimagining work, workers, and workplaces for the future," Global Business and Organisational Excellence, 2023.

[5] L. Fortunati and A. Edwards, "Framing the psycho-social and cultural aspects of human-machine communication," Human-Machine Communication, 2022.

[6] E. K. Chiou and J. D. Lee, "Trusting Automation: Designing for Responsivity and Resilience," Human Factors, 2023.

[7] T. Samad, "Human-in-the-Loop Control and Cyber–Physical–Human Systems: Applications and Categorization," in Cyber–Physical–Human Systems: Fundamentals and Applications, A. M. Annaswamy, et al., Eds. Hoboken, NJ: John Wiley & Sons, Inc., 2023.

[8] EC JRC, “Research and Innovation in Transport Safety and Resilience in Europe”, 2023.

[9] LABYRINTH Project (H2020 - Grant ID No 861696)

[10] SAFEDRONE Project (SESAR 2020 - Grant ID No 783211)

[11] [10] MoNIfly Project (H2020 - Grant ID: No. 723509)

[12] COMP4DRONES Project (ECSEL JU Grant ID No 826610)

[13] TINDAIR Project (SESAR 2020 - Grant ID No 101017677)

[14] AURORA Project (SESAR 2020 - Grant ID No 699340)

[15] GAUSS Project (H2020 - Grant ID: No. 101004255)

[16] SAFIR-Med Project (SESAR 2020 - Grant ID No 101017701)

[17] ASSURED-UAM Project (H2020 - Grant ID: No. 101006696)

[18] MAHALO Project (SESAR 2020 - Grant ID No 892970)

[19] PJ16 CWP HMI Project (SESAR 2020 - Grant ID No 734141)

[20] FARO Project (SESAR 2020 - Grant ID No 892542)

[21] JARVIS Project (DES - Grant ID No 101114692)

[22] DARWIN Project (DES - Grant ID No 101114733)

[23] LOKI - Collaboration of aviation operators and AI systems (DLR)

[24] SafeClouds Project (H2020 - Grant ID: No. 653884)

[25] PJ11 CAPITO Project (SESAR 2020 - Grant ID No 732996)

[26] ODESSA Project (H2020 - Grant ID: No. 821263)

[27] SENSORIANCE Project (H2020 - Grant ID: No. 785315)

[28] WINFC Project (FP7 - Grant ID: No. 632536)

[29] VISION Project (H2020 - Grant ID: No. 690811)

[30] STRESS Project (H2020 - Grant ID: No. 699381)

[31] REPS Project (H2020 - Grant ID: No. 101008005)

[32] SafeNcy Project (H2020 - Grant ID: No. 864771)

[33] SAFELAND Project (SESAR 2020 - Grant ID No 890599)

[34] GRIMASSE Project (SESAR 2020 - Grant ID No 776379)

[35] HAIKU Project (HE - Grant ID No 101075332)

[36] SAE ARP4761 – Guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment

[37] EUROCAE ED-135 – Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment

[38] RTCA DO-178C/EUROCAE ED-12C: Software Considerations in Airborne Systems and Equipment Certification

[39] EASA, “Concept Paper: Guidance for Level 1&2 machine learning applications”, Issue 02, March 2024.

[40] HLEG-AI, “Ethics guidelines for trustworthy AI”, 8 April 2019.

[41] HLEG-AI (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. 17 July 2020.

[42] To monitor the progress of this initiative, please, see: https://eurocae.net/news/posts/2019/june/new-working-group-wg-114-artificial-intelligence/

[43] ISO/IEC TR 5469:2024 - Artificial Intelligence - Functional safety and AI systems

[44] ISO/IEC CD TS 6254 - Information technology – Artificial intelligence – Objectives and approaches for explainability and interpretability of ML models and AI systems

[45] ISO/IEC 25000 - Software Quality Requirements and Evaluation

[46] To monitor the progress of this initiative, please, see: https://engagestandards.ieee.org/ieeecertifaied.html

[47] SESAR JU, Multiannual Work Programme – 2022-2031, 2021.

[48] SESAR JU (2019, December 17). European ATM Master Plan 2020.

[49] Decision (EU) 2022/2481 – Digital Decade Policy Programme 2030.

[50] EC. Artificial Intelligence for Europe (COM/2018/237 final), 2018.

[51] SESAR JU, “European ATM Master Plan Stakeholder consultation workshop pre-read material”, 22-23 April 2024