

# APPLYING NATURAL LANGUAGE PROCESSING IN INNOVATION ECONOMICS

*Dr. Daniel Weiss,  
Research Group Leader „Market analysis, -design and innovation strategies“,  
German Aerospace Center (DLR), Institute for Transport Research*

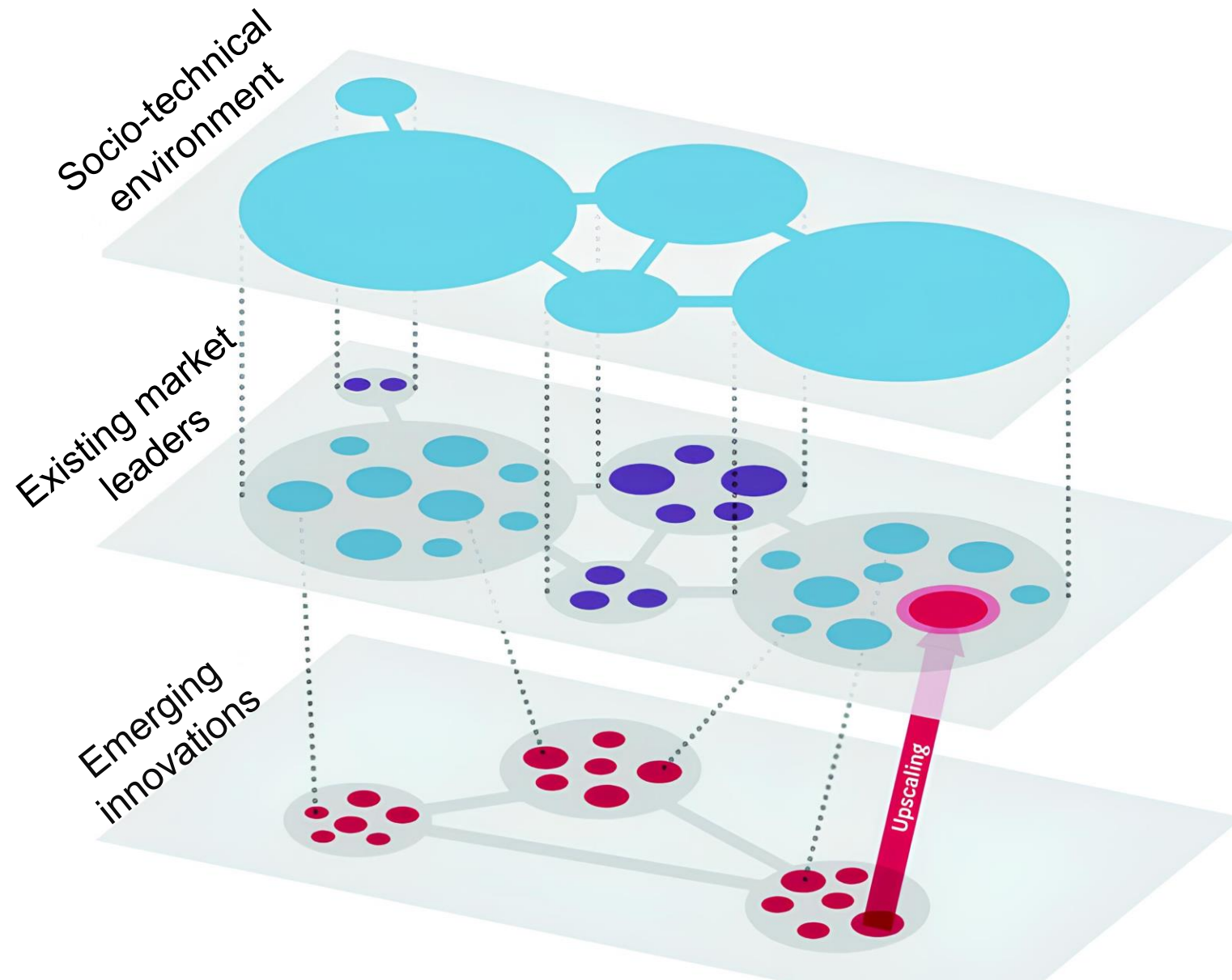


- Why text analysis
- Applications to mobility
- Translation into decision-making
- Discussion & Outlook

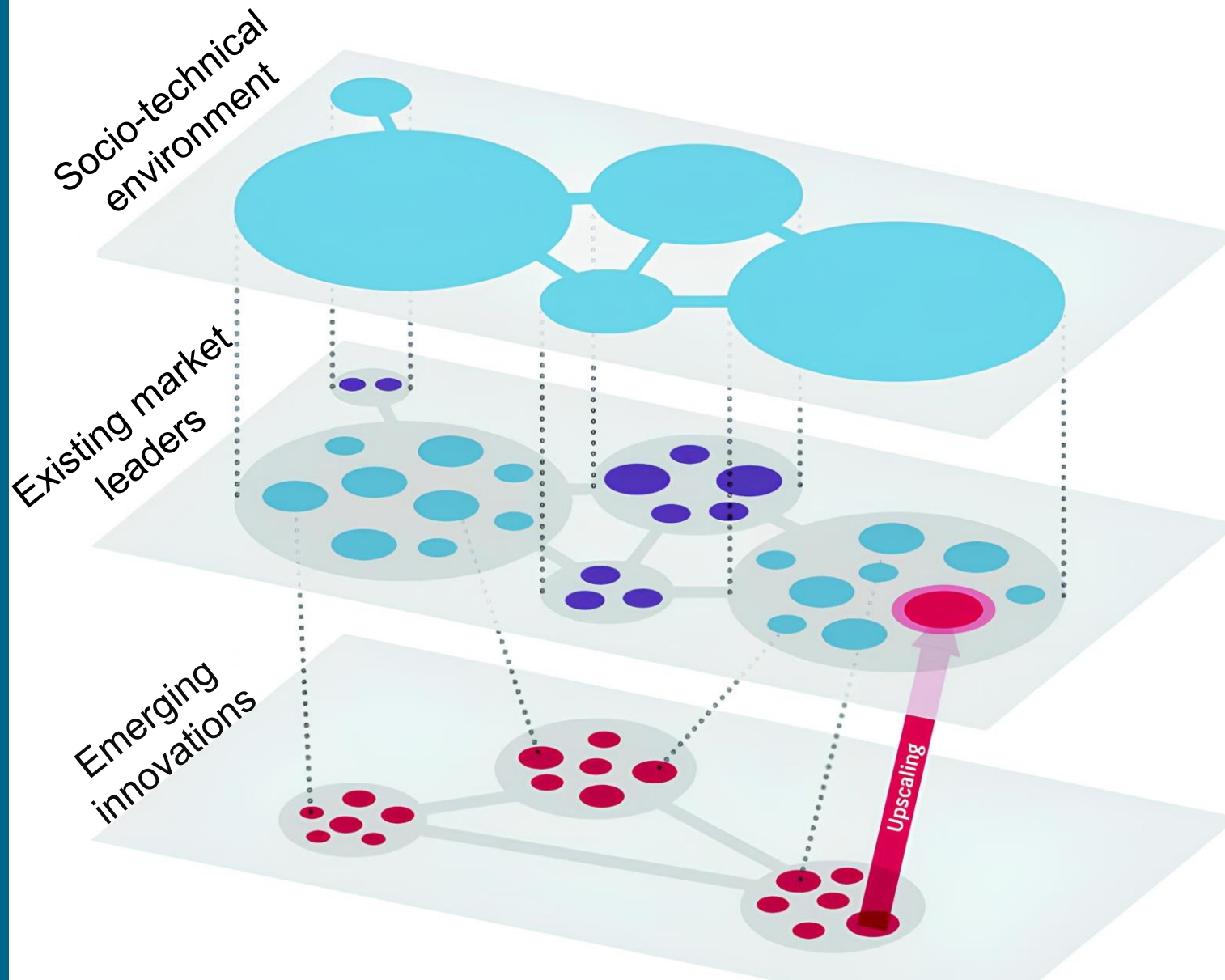


# WHY TEXT ANALYSIS

# Managing Complex Transition Processes



# Managing Complex Transition Processes



Currently:

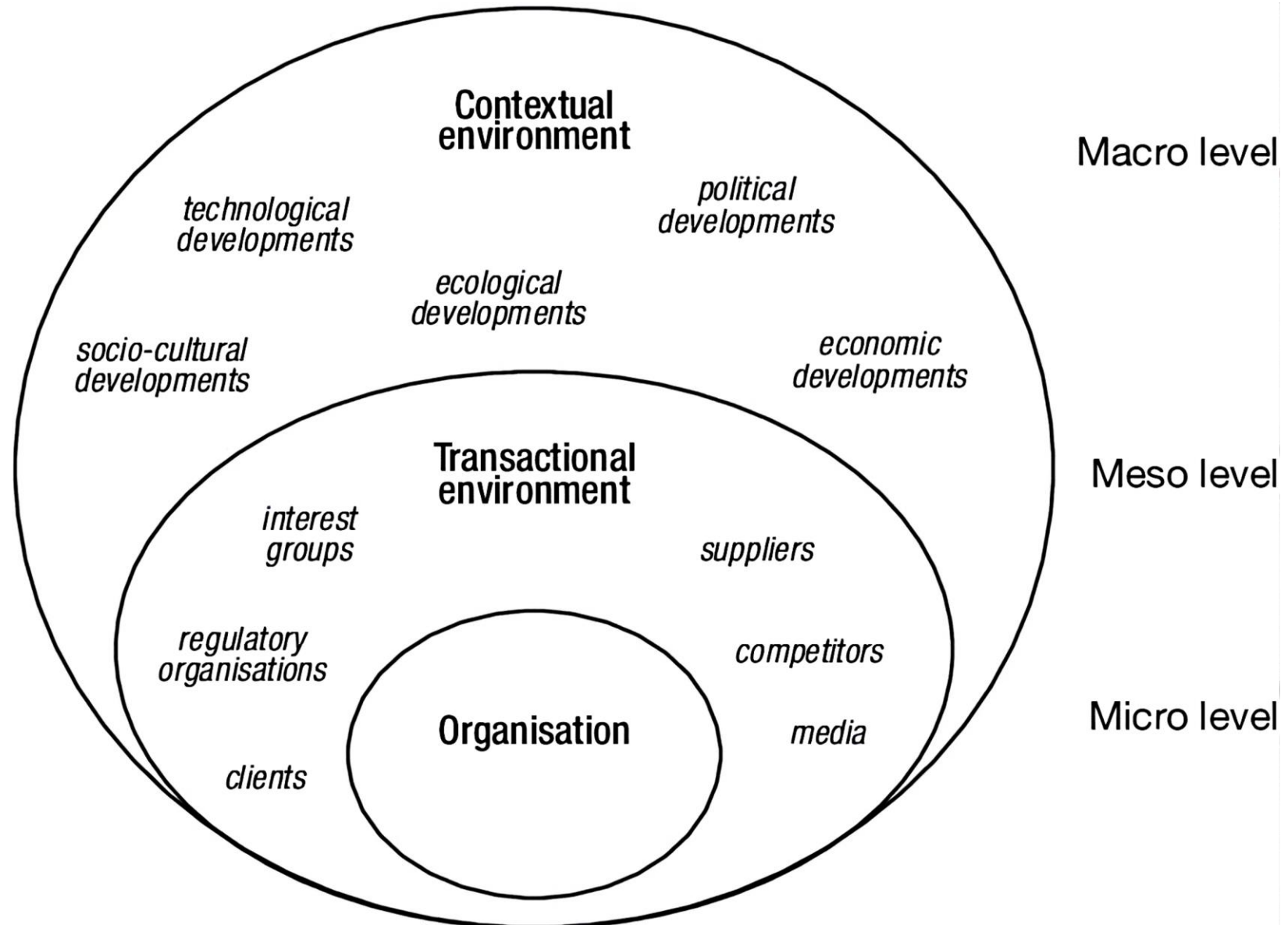
- Primarily qualitative methods
- Only individual indicators for technology development

We need:

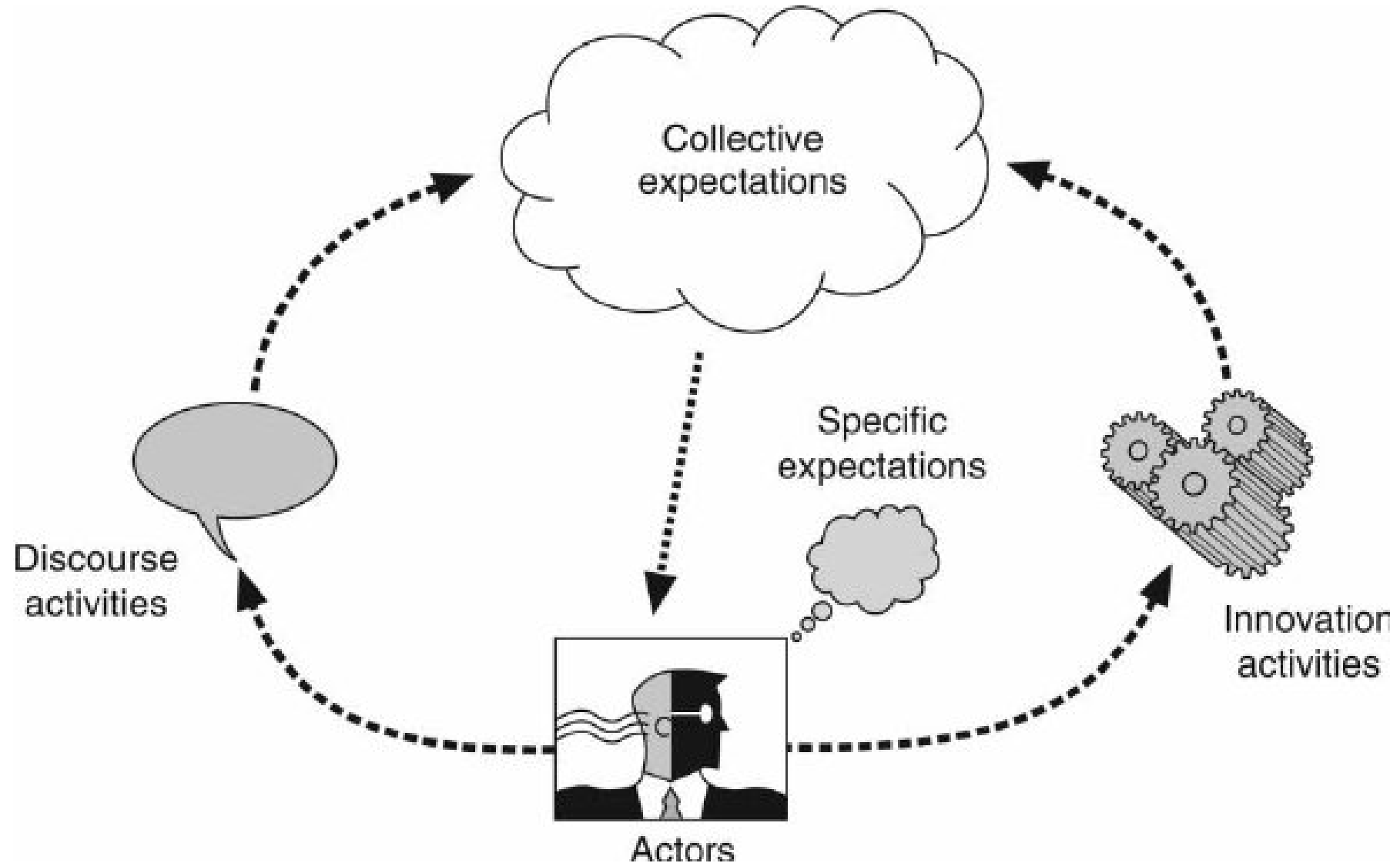
- Tools for ad-hoc analyses and holistic monitoring are missing
- Up to 90% of all data is unstructured text data (Berger et al. 2019)



# How to navigate complex transition processes?



# Discourse Dynamics and Text Data

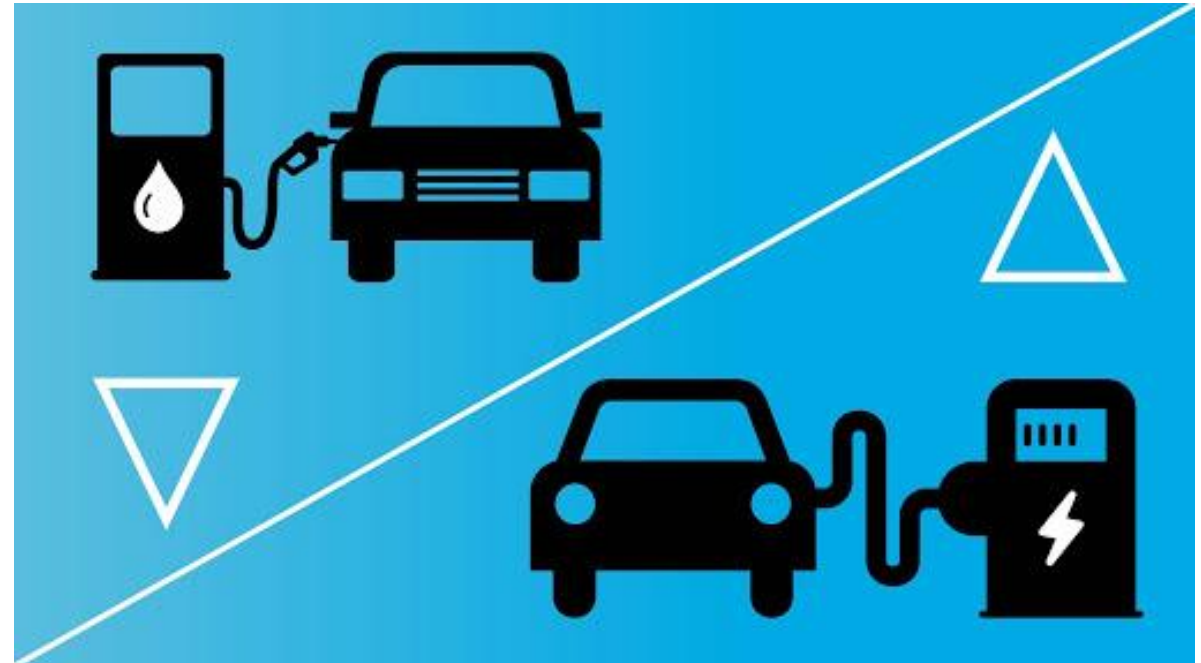


A TEXT-BASED MONITORING TOOL FOR THE LEGITIMACY AND GUIDANCE OF TECHNOLOGICAL INNOVATION SYSTEMS (WEISS AND NEMECZEK 2021)

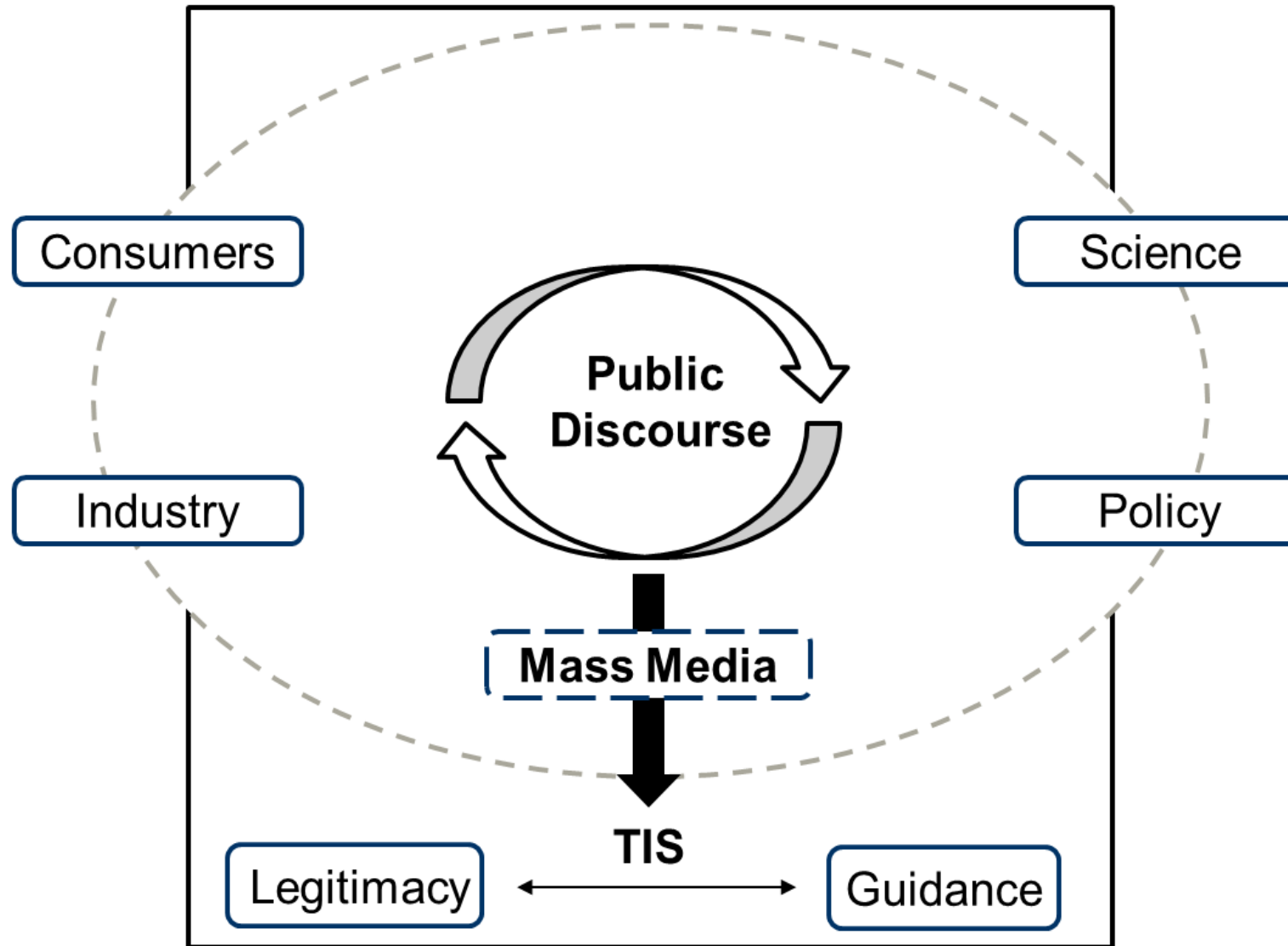


# Liability of newness and discourse activities

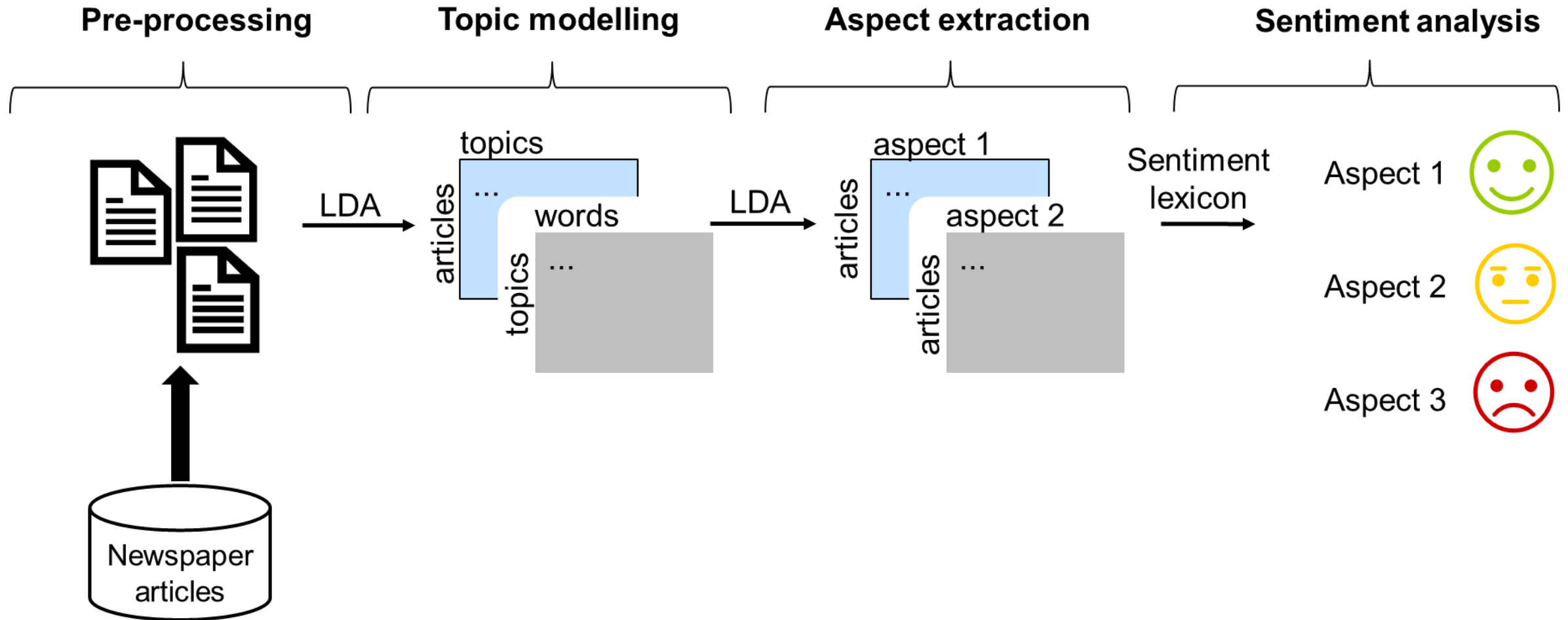
- ‘Liability of newness’ of BEVs
  - Compliance issues and uncertainty (*Legitimacy/Guidance*)
  - Reflected in public discourse but drawbacks of manual text analysis
- NLP-based monitoring tool



# Public discourse – Technology Legitimacy, and Guidance



# Processing pipeline



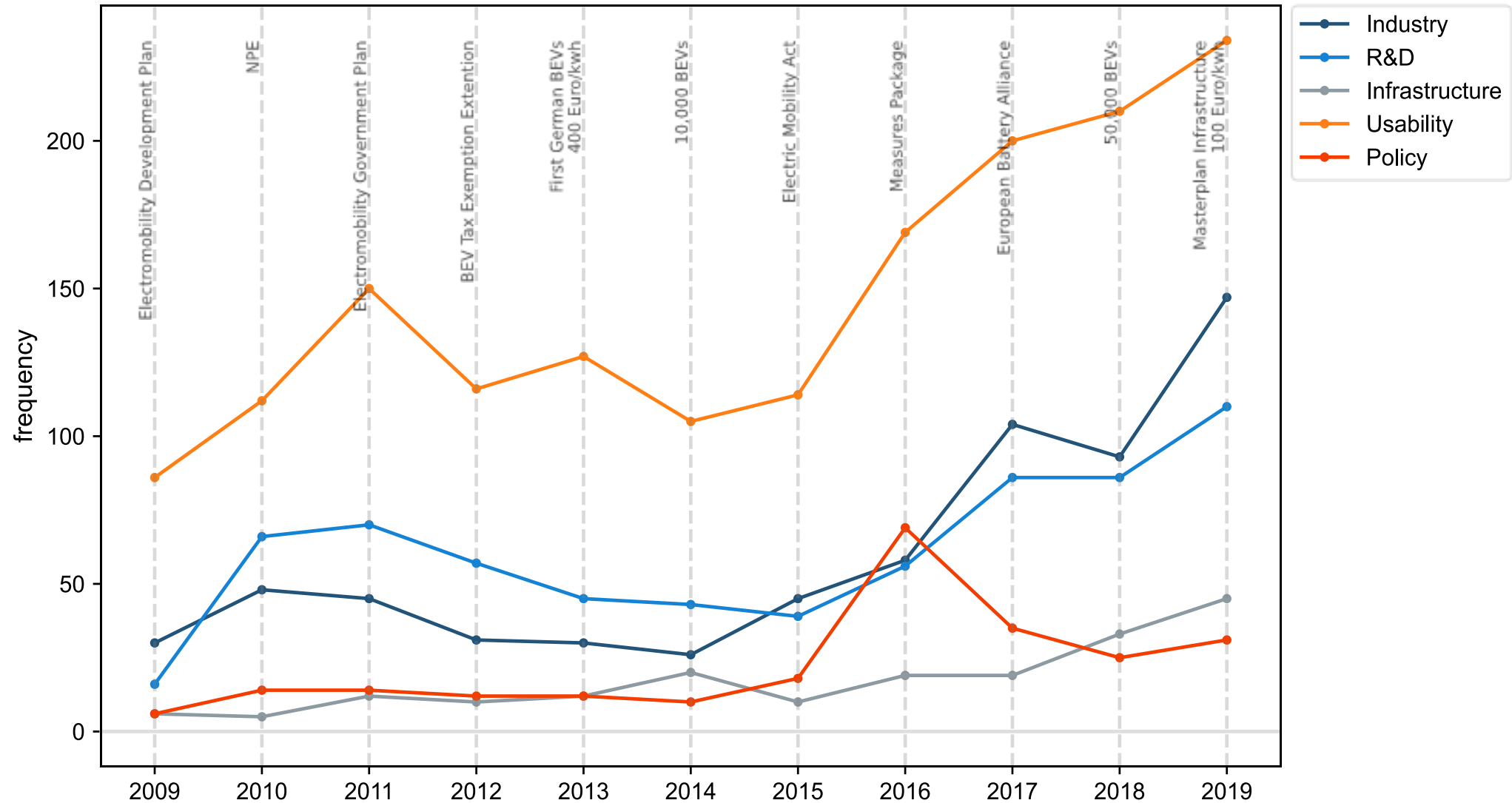


# LDA topic keywords and probabilities for BEV

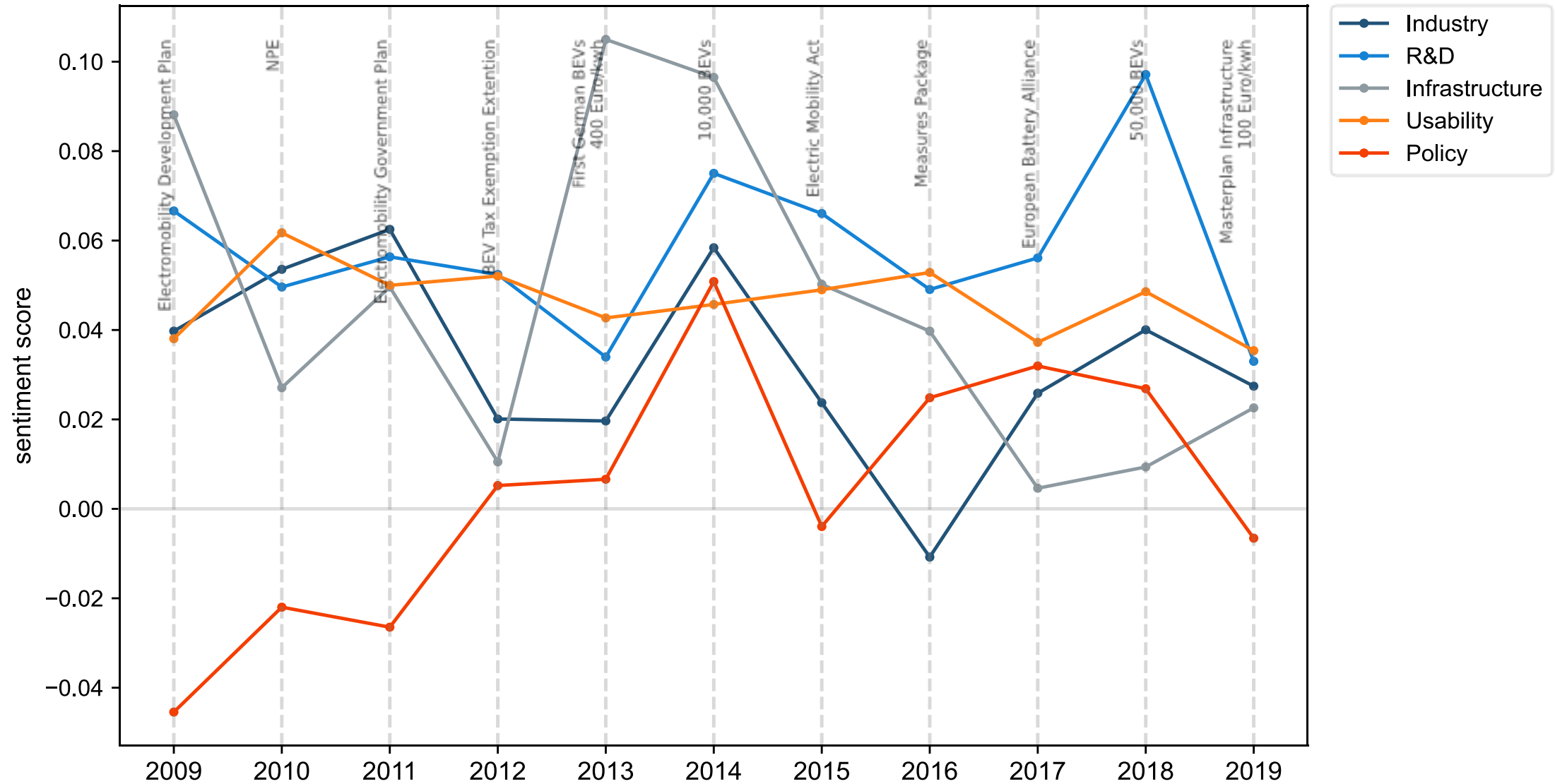


Topic	Distribution of the most important topic keywords
<b>Industry</b>	0.005*"milliarde" + 0.004*"bus" + 0.003*"konzern" + 0.003*"produktion" + 0.003*"kamenz" + 0.003*"mitarbeiter" + 0.003*"elektrobus" + 0.003*"produzieren" + 0.002*"projekt" + 0.002*"verkaufen" + 0.002*"batteriezelle" + 0.002*"branche" + 0.002*"elektromotor" + 0.002*"werk" + 0.002*"bereich" + 0.002*"industrie" + 0.002*"zulieferer" + 0.002*"firma" + 0.002*"investieren" + 0.002*"geschäft"
<b>R&amp;D</b>	0.004*"e-bike" + 0.003*"team" + 0.003*"projekt" + 0.003*"firma" + 0.003*"motor" + 0.003*"idee" + 0.003*"einsatz" + 0.003*"professor" + 0.003*"technologie" + 0.003*"akku" + 0.003*"rennen" + 0.002*"material" + 0.002*"thema" + 0.002*"meter" + 0.002*"welt" + 0.002*"institut" + 0.002*"mobilität" + 0.002*"lithium" + 0.002*"hochschule" + 0.002*"fahrrad"
<b>Infrastructure</b>	0.005*"ladestation" + 0.005*"stadtwerke" + 0.004*"standort" + 0.004*"ladesäule" + 0.003*"akku" + 0.003*"station" + 0.003*"wagen" + 0.003*"laden" + 0.003*"tesla" + 0.003*"meter" + 0.003*"anlage" + 0.003*"netz" + 0.003*"minute" + 0.003*"fahrer" + 0.003*"steckdose" + 0.003*"fahrt" + 0.003*"dauern" + 0.002*"kilometerprostunde" + 0.002*"leistung" + 0.002*"monat"
<b>Usability</b>	0.006*"tesla" + 0.004*"wagen" + 0.003*"ampera" + 0.003*"akku" + 0.003*"idrei" + 0.003*"leaf" + 0.003*"sekunde" + 0.003*"model" + 0.003*"fahrer" + 0.003*"steckdose" + 0.003*"laden" + 0.002*"dollar" + 0.002*"ladestation" + 0.002*"haus" + 0.002*"pferdestärke" + 0.002*"meter" + 0.002*"netz" + 0.002*"stromer" + 0.002*"hören" + 0.002*"technik"
<b>Policy</b>	0.006*"prämie" + 0.005*"antrag" + 0.005*"kaufprämie" + 0.004*"zelle" + 0.004*"diesel" + 0.004*"kaufen" + 0.003*"staat" + 0.003*"bund" + 0.003*"beantragen" + 0.003*"verein" + 0.003*"verbrennungsmotor" + 0.003*"förderung" + 0.003*"milliarde" + 0.003*"wirtschaft" + 0.003*"bundesregierung" + 0.003*"firma" + 0.003*"bundesamt" + 0.003*"regierung" + 0.003*"rechnen" + 0.003*"verbraucher"

# Media attention per topic for BEVs in Germany

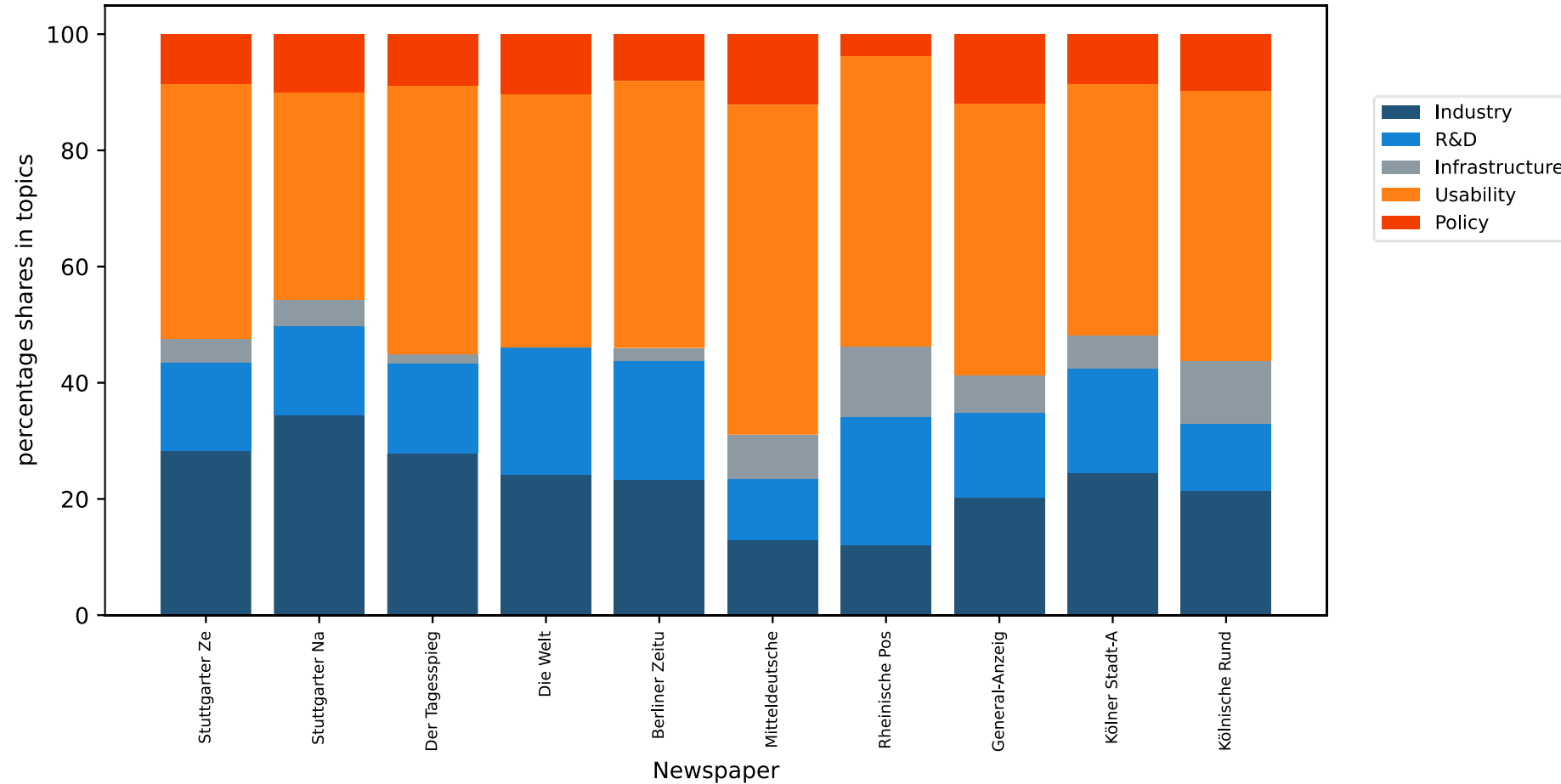


# Media sentiment per topic for BEVs in Germany

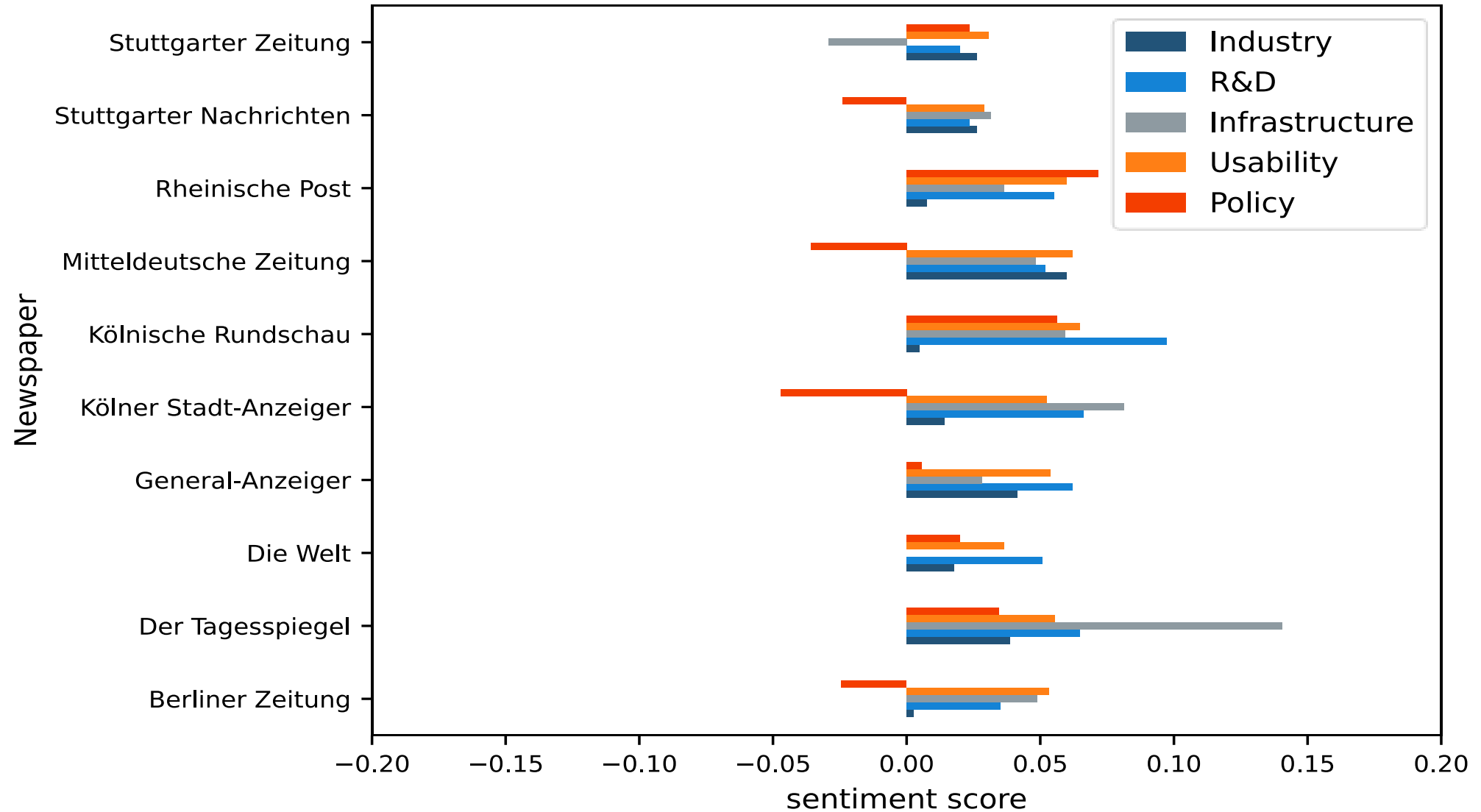




# Differences across regional newspapers



# Differences across regional newspapers



**A MEDIA-BASED INNOVATION INDICATOR: EXAMINING DECLINING  
TECHNOLOGICAL INNOVATION SYSTEMS (WEISS AND NEMECZEK 2022)**



# Measuring the decline of a technology

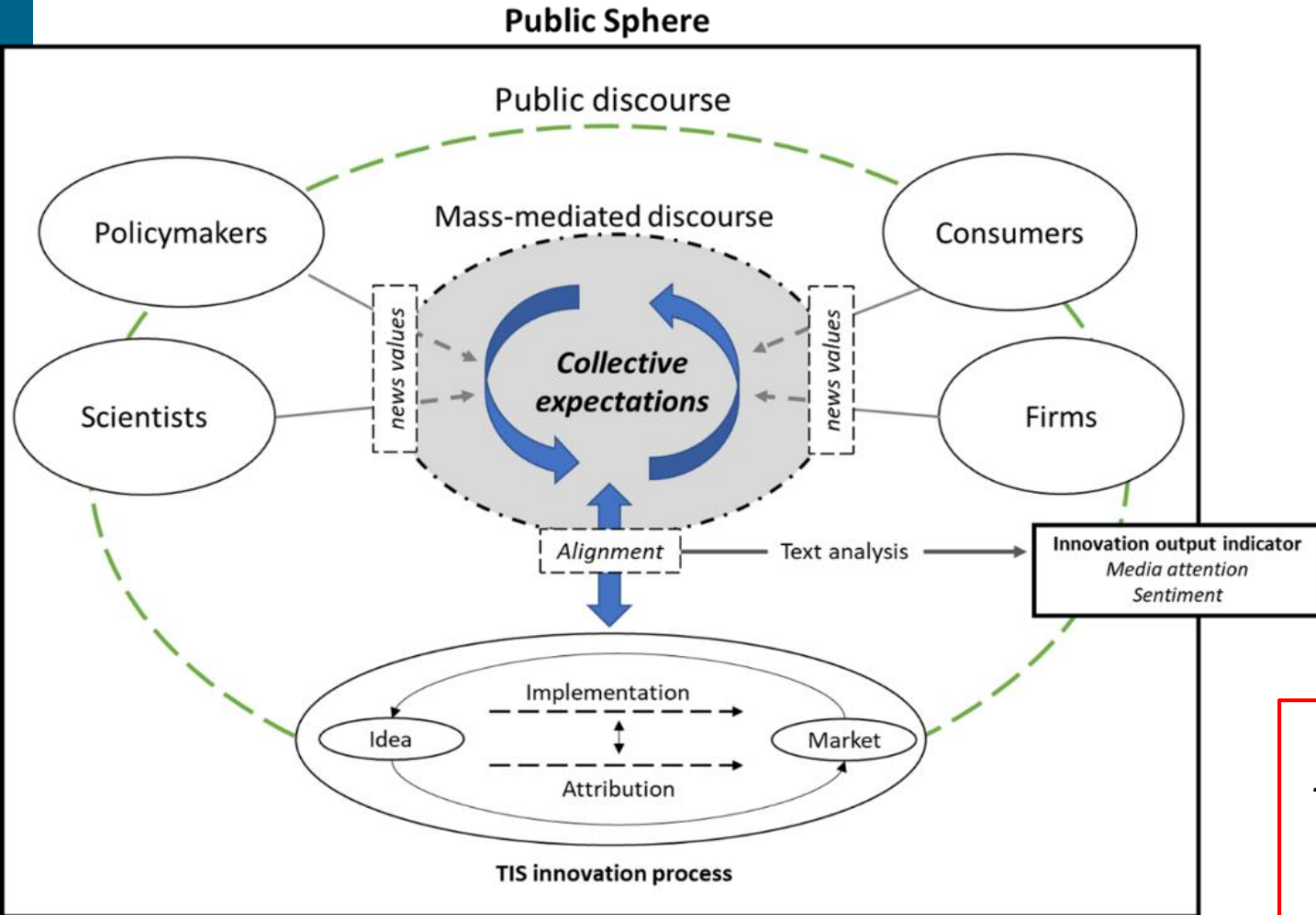


- Shortcomings of current technology decline indicators such as production, sales and patent data (Isoaho and Markard 2020, Markard et al. 2020, Weiss and Scherer 2021)
  - Delayed and missing granularity (Kinne and Axenbeck 2020)
  - Limitations of patent data (Eggink 2012)
  - Neglecting collective expectations and preferences (van Lente and Rip 1998, Brown and Michael 2003, Borup et al. 2006)

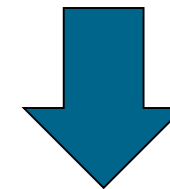


- Novel mass-media-based innovation output indicator illustrated with ICE
- Methodological framework based on NLP (Bellstam et al. 2021, Weiss and Nemeček 2021)

# Conceptual foundations



- Aggregated mass-mediated discourse as a proxy of collective expectations and preferences (Konrad et al. 2012, Waldherr 2012, Dehler-Holland et al. 2021, Weiss and Nemeček 2021)
- Technology-specific news values guide media attention for innovations (Waldherr 2008, 2012)
  - Relevance, timeliness, and relatedness to prevailing societal problems

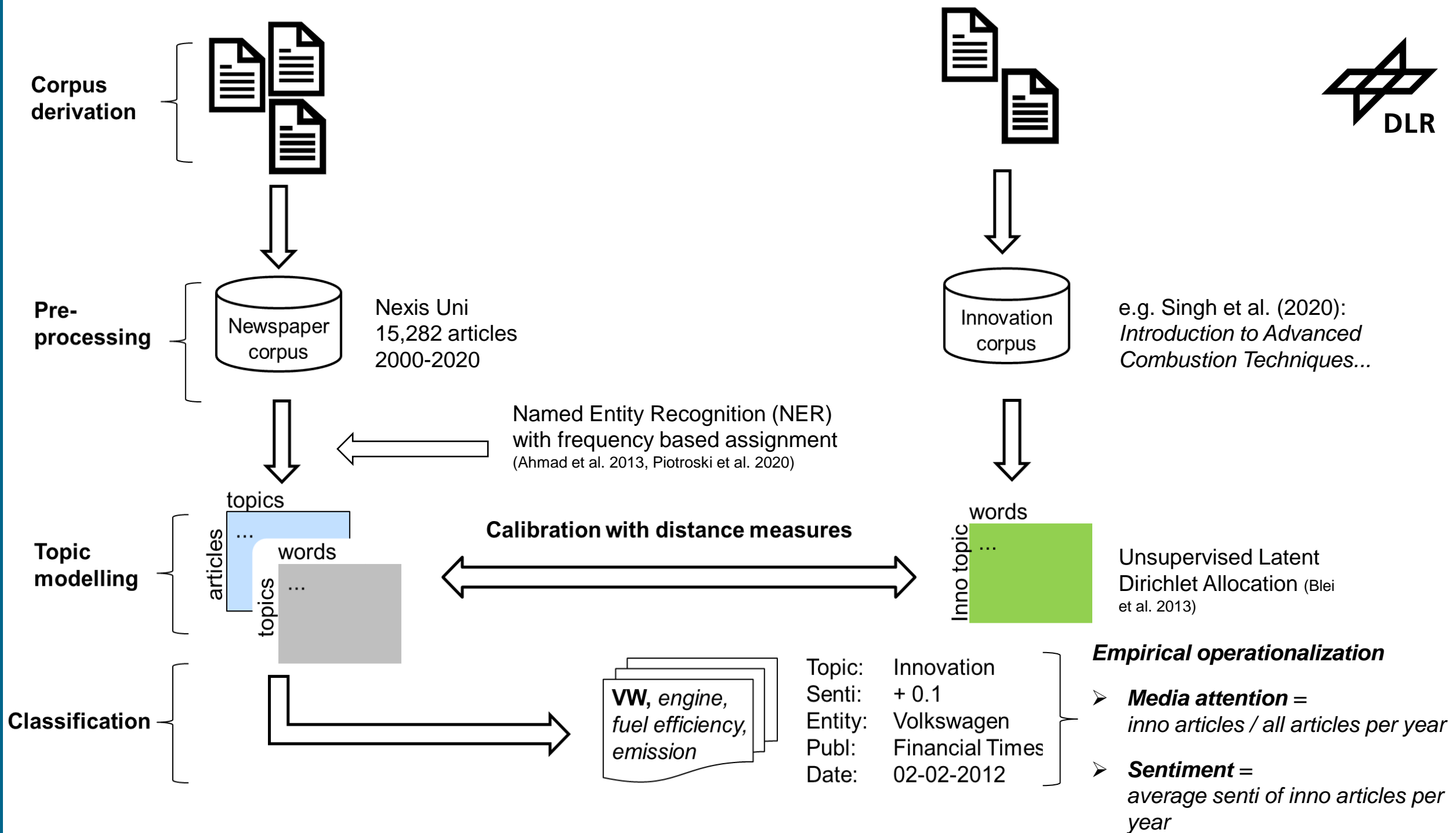


*Assumption:*

*The media reports on innovations that are aligned with collective expectations and preferences*

Media attention + sentiment indicates demand and support for the innovation

Source: own depiction with reference to Waldherr (2008, 2012)

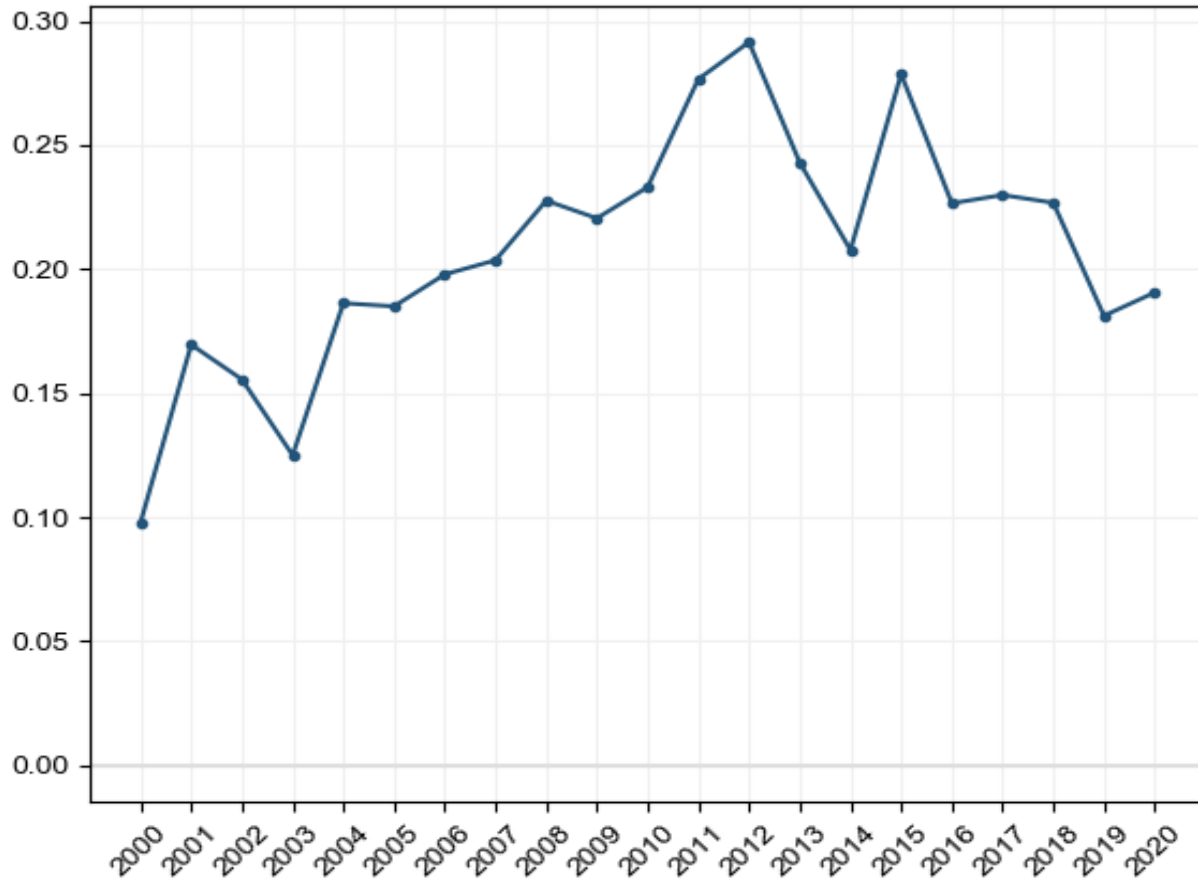




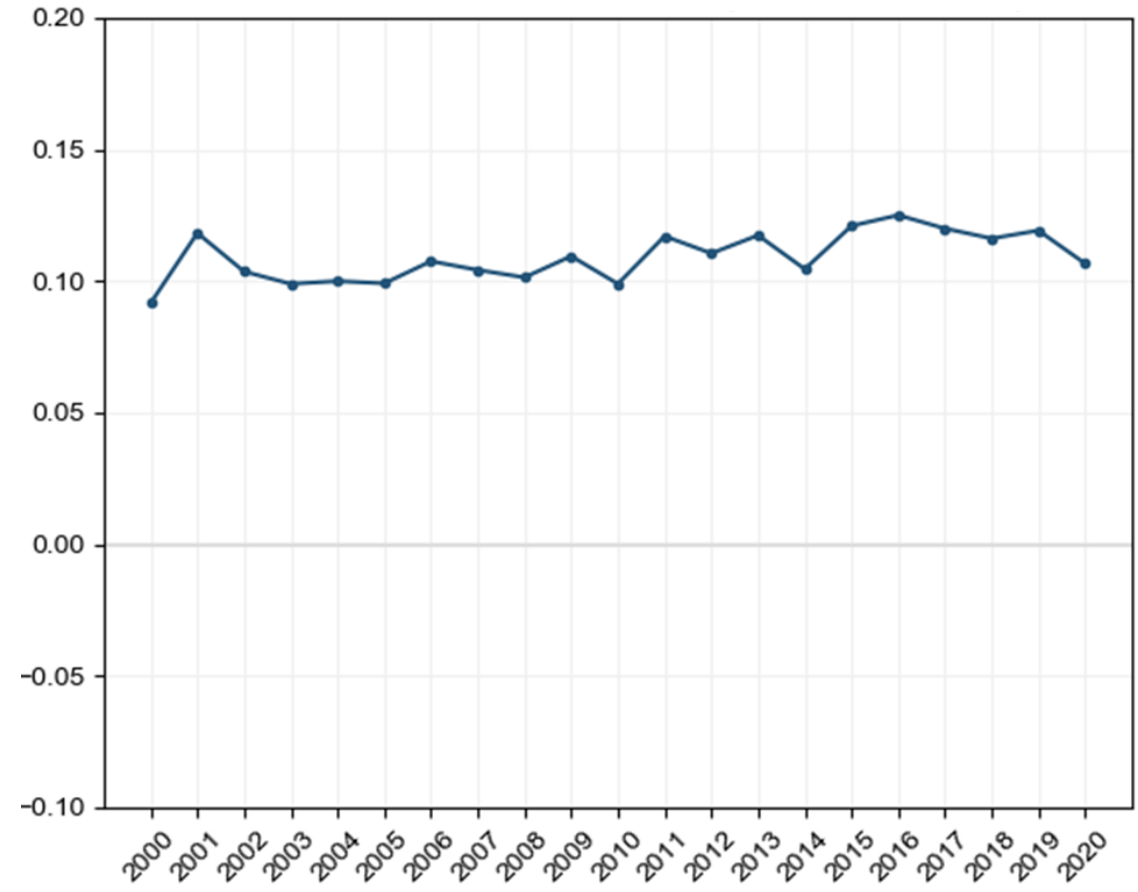


# Indicator results for ICE, English newspapers 2000-2020

## Share of innovation articles



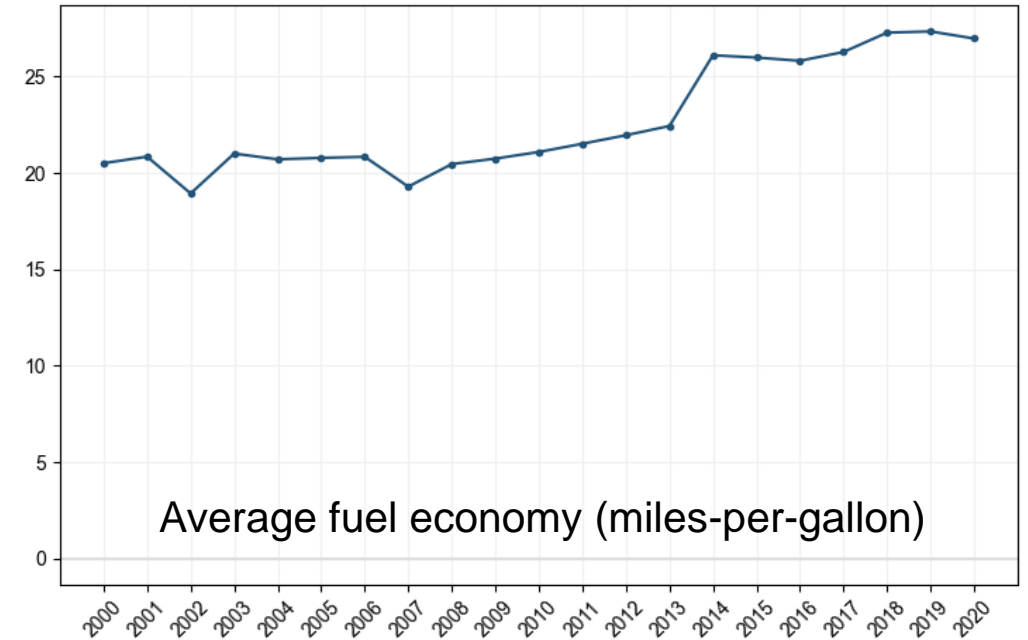
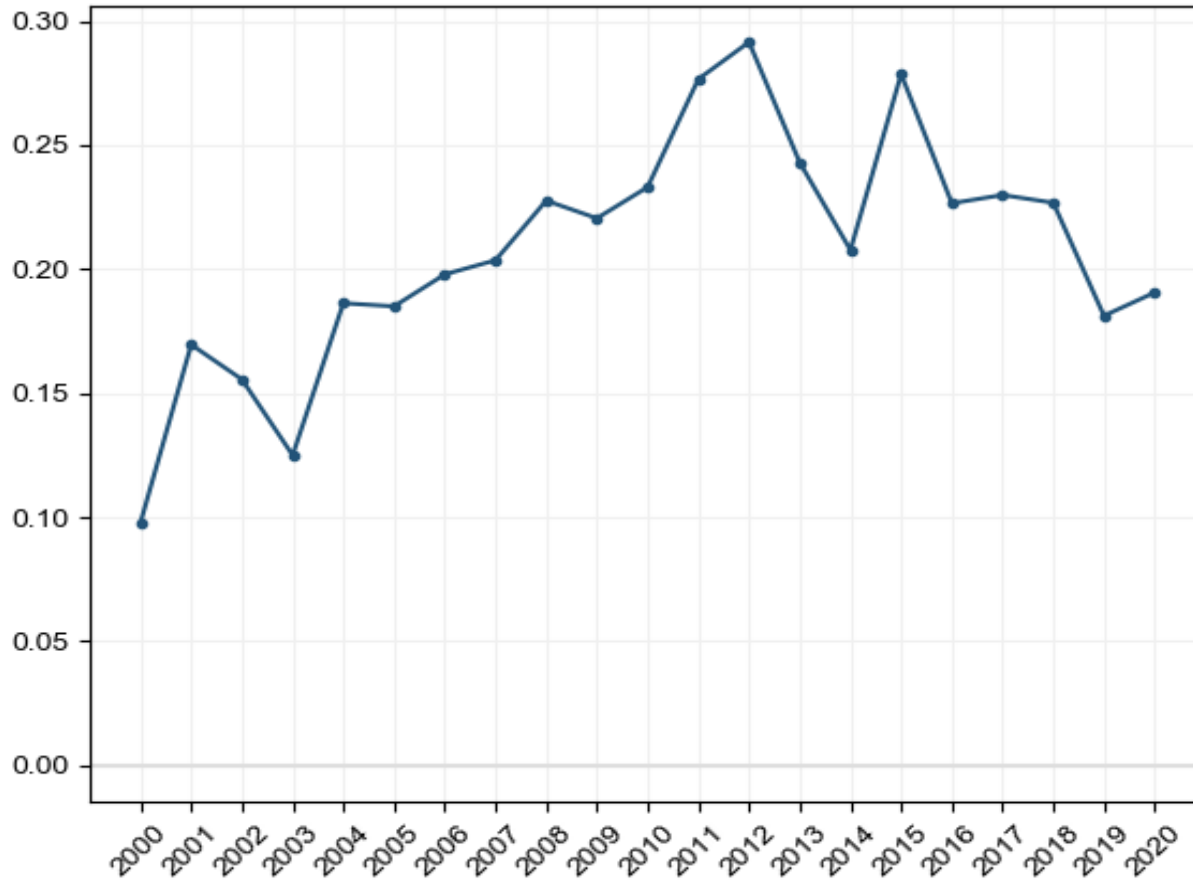
## Sentiment



- **Decreasing trend after 2015, although positive sentiment. Robust results.**

# Indicator results for ICE, English newspapers 2000-2020

## Share of innovation articles

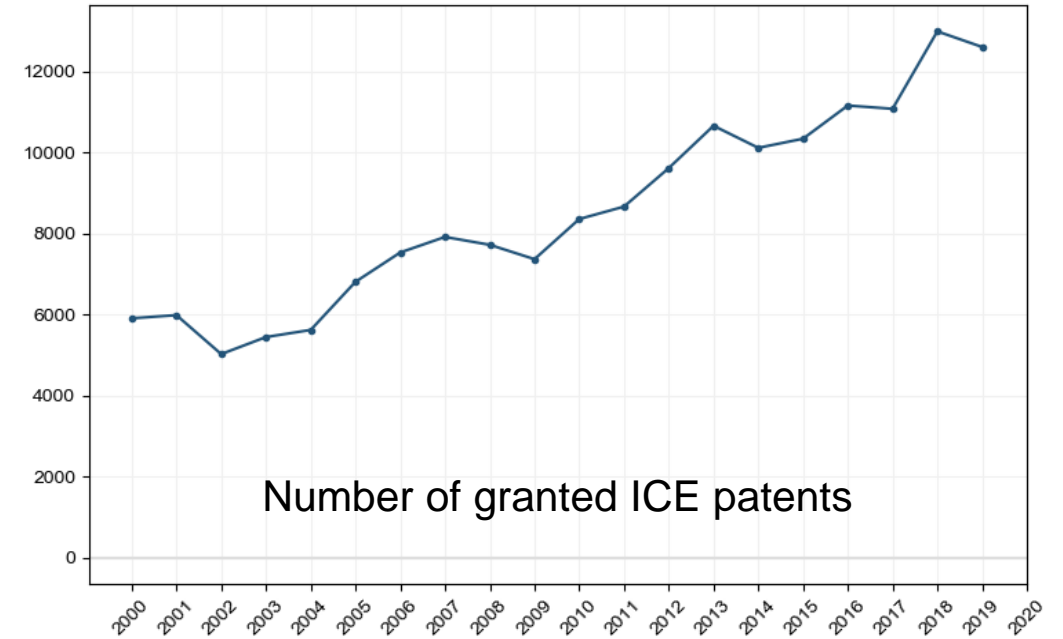
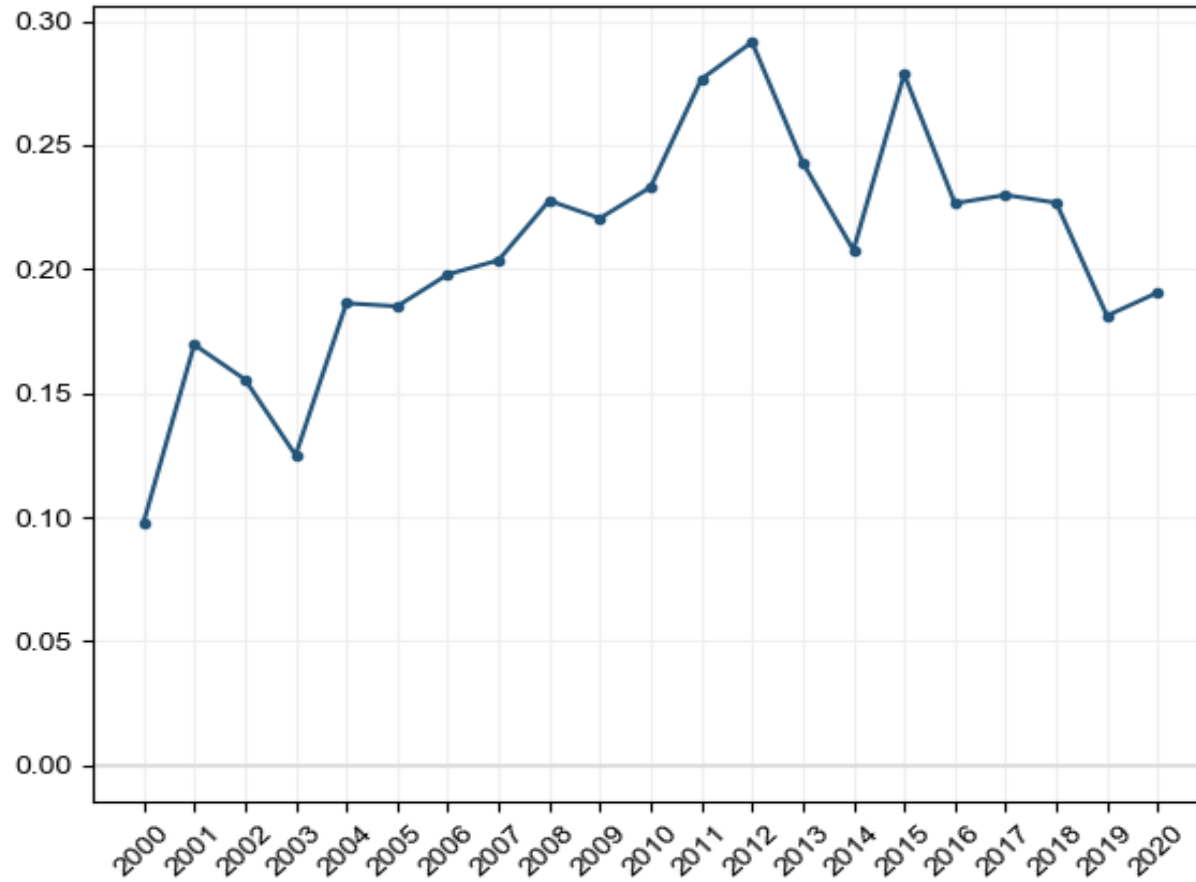


Source: EPA 2020 (filtered by cylinder count, model year average)

- **Decreasing trend after 2015, although positive sentiment. Robust results.**

# Indicator results for ICE, English newspapers 2000-2020

## Share of innovation articles

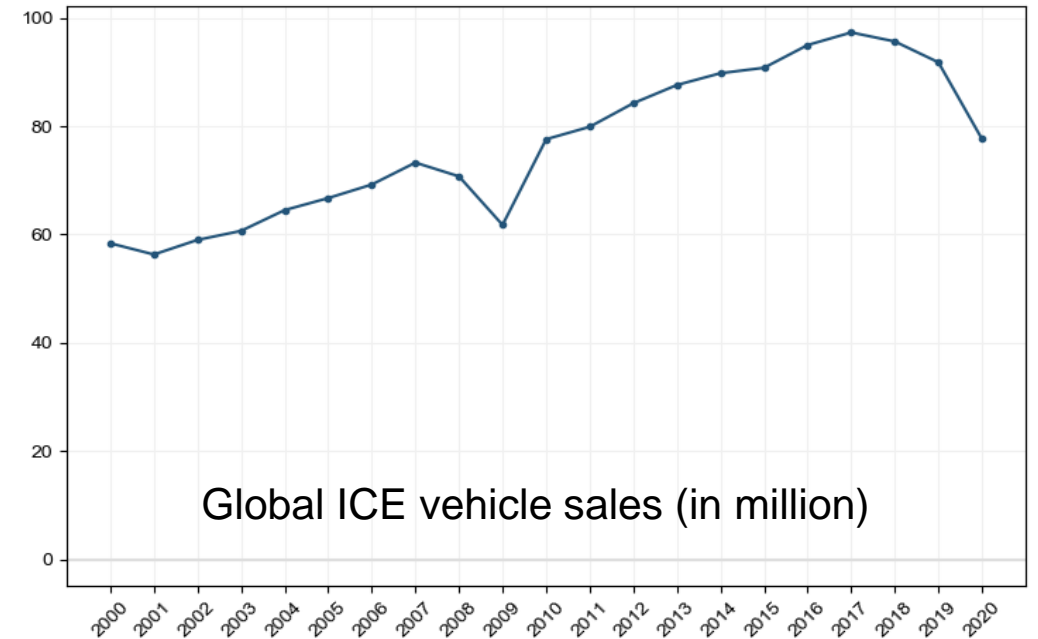
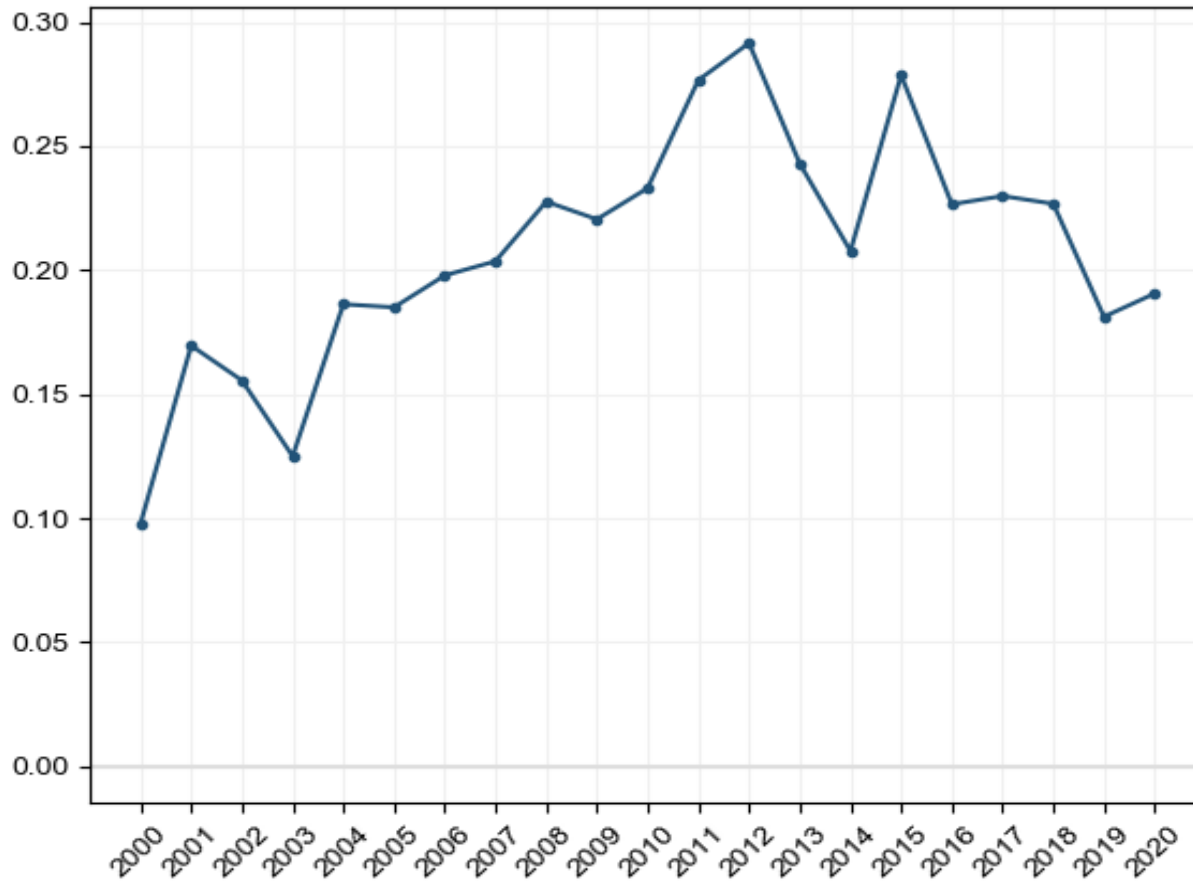


Source: EPO (with IPC codes from Aghion et al. (2016))

- **Decreasing trend after 2015, although positive sentiment. Robust results.**

# Indicator results for ICE, English newspapers 2000-2020

## Share of innovation articles



Source: IEA 2021

- **Decreasing trend after 2015, although positive sentiment. Robust results.**

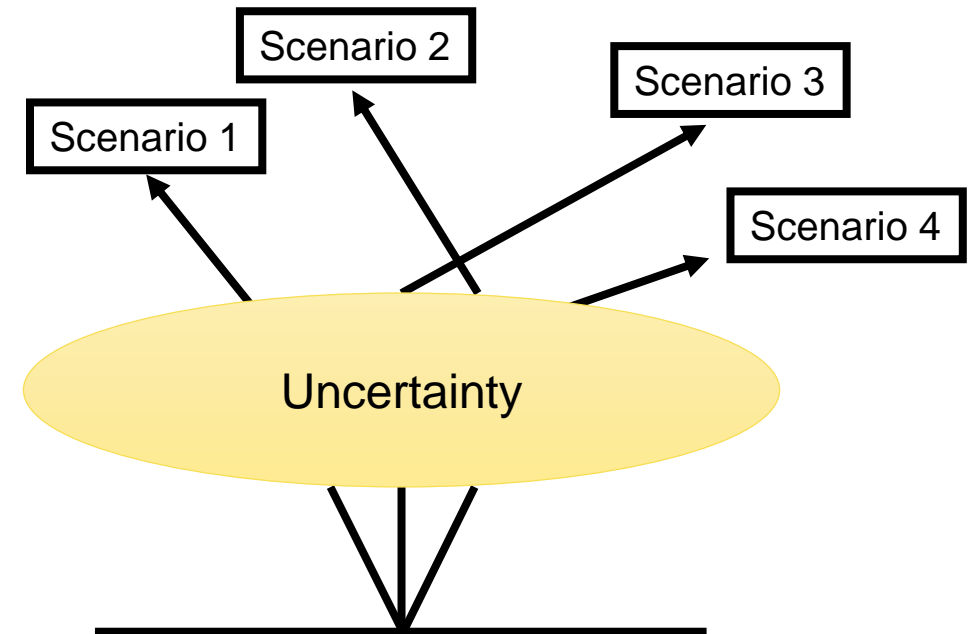


## Indicator results for ICE, English newspapers 2000-2020

<b>Firm</b>	<b>Total number of innovation articles</b>	<b>Average sentiment</b>
<b>Mazda</b>	53	0,17
<b>Bmw</b>	50	0,16
<b>Daimler</b>	33	0,17
<b>Wartsila</b>	29	0,14
<b>Honda</b>	28	0,16
<b>Infiniti</b>	28	0,21
<b>Ford</b>	26	0,22
<b>Peugeot</b>	19	0,21
<b>Achates</b>	17	0,15
<b>Hyundai</b>	16	0,17

# Further opportunities for text analysis

- Map the whole innovation system
- Use foresight techniques
- LLMs for ,hybrid‘ approaches
- Dedicated evaluation frameworks



# TRANSLATION INTO DECISION-MAKING

# Complexity and 'noise' hinder translation into decision-making



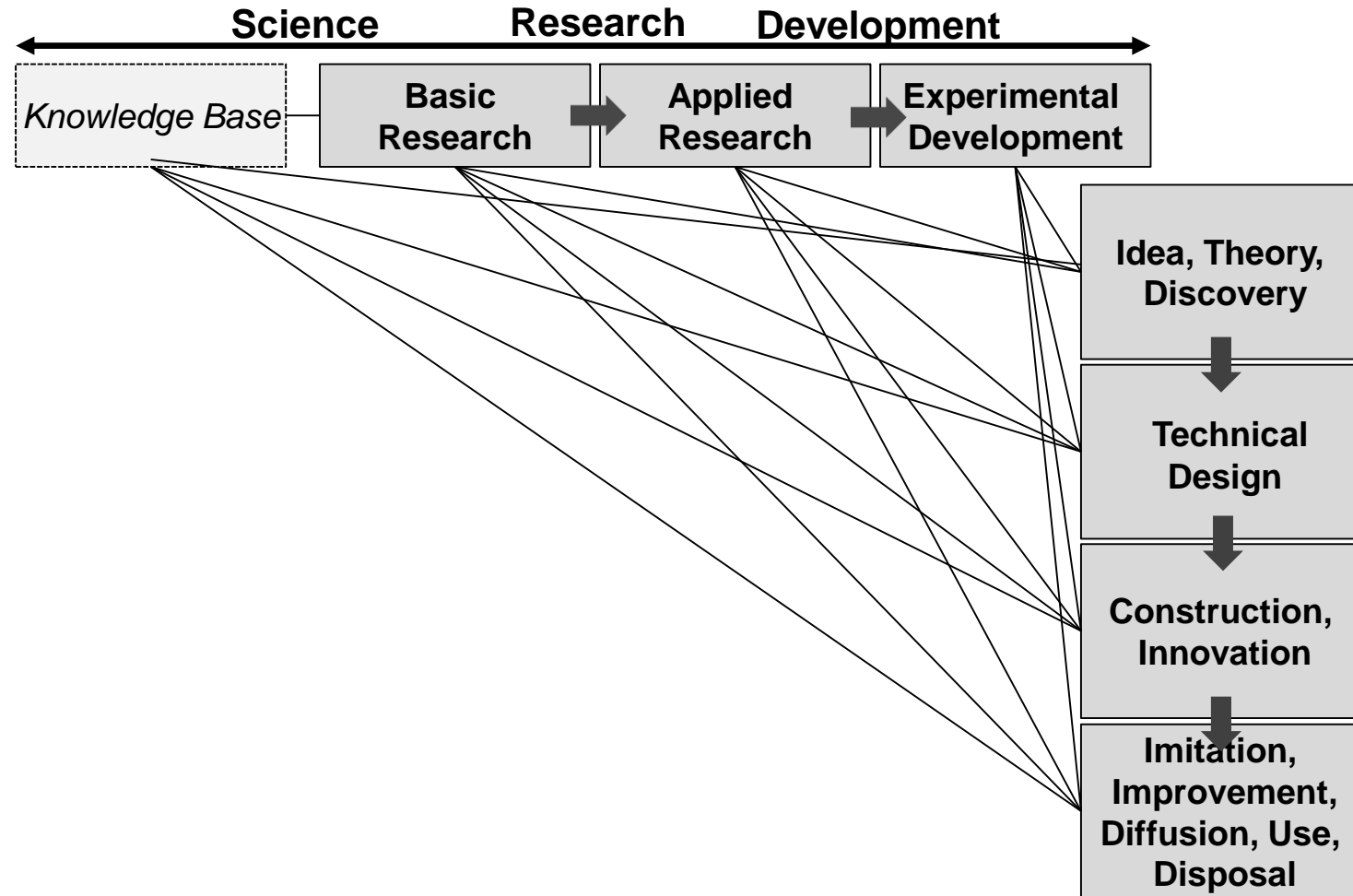
## ➤ Some key issues of previous approaches:

- Biases in data-driven prioritization of impact factors
- Superficial frameworks, such as PESTEL (Political, Economic, Social, Technological, Environmental, Legal)



- Utilize the explanatory power of innovation economics
  - Technological Innovation Systems (Hekkert et al. 2007, Markard 2020)
  - Multi-Level-Perspective (Geels 2002)

# Measuring Technological Progress using Innovation Indicators

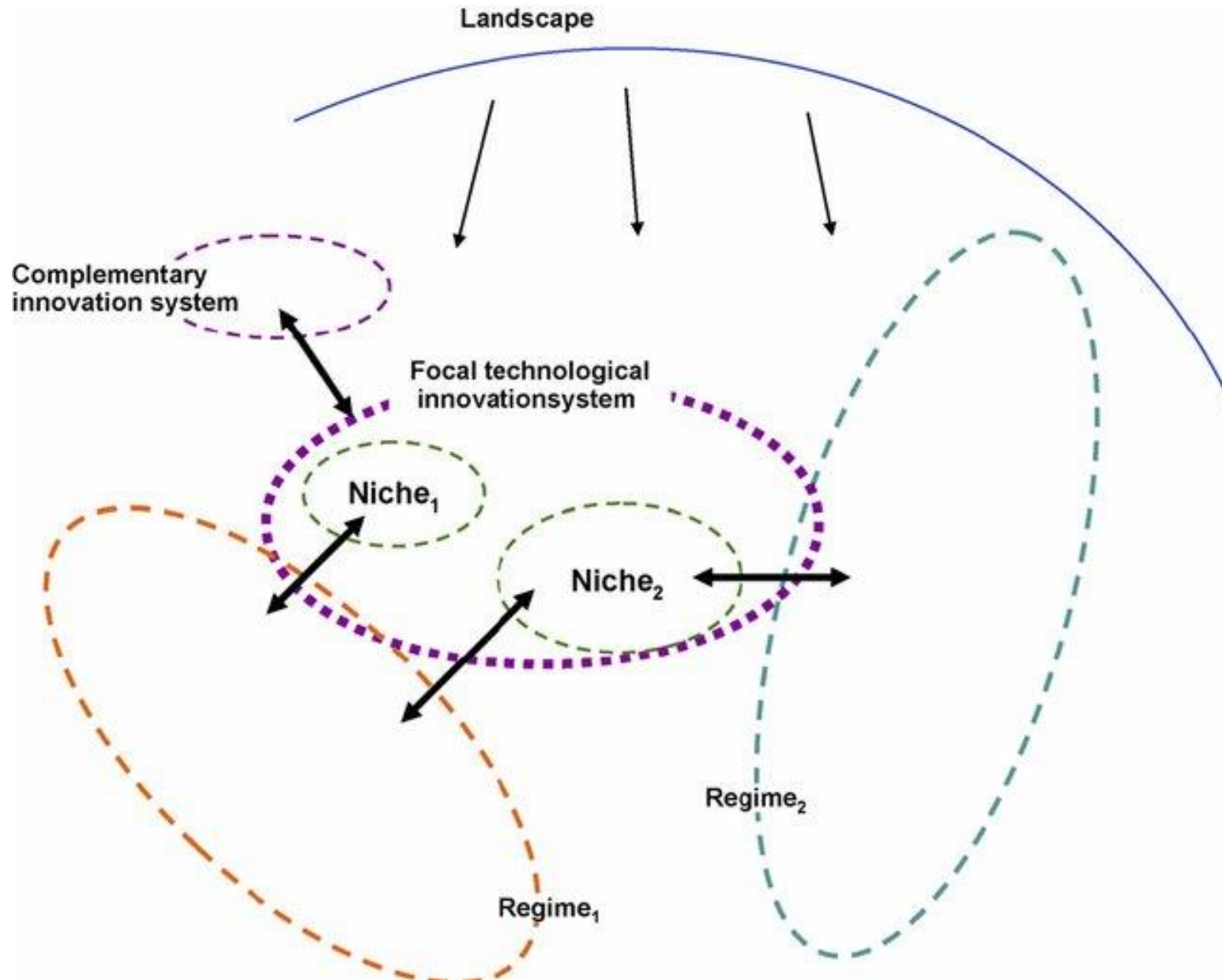


# Innovation indicators using text data - Examples



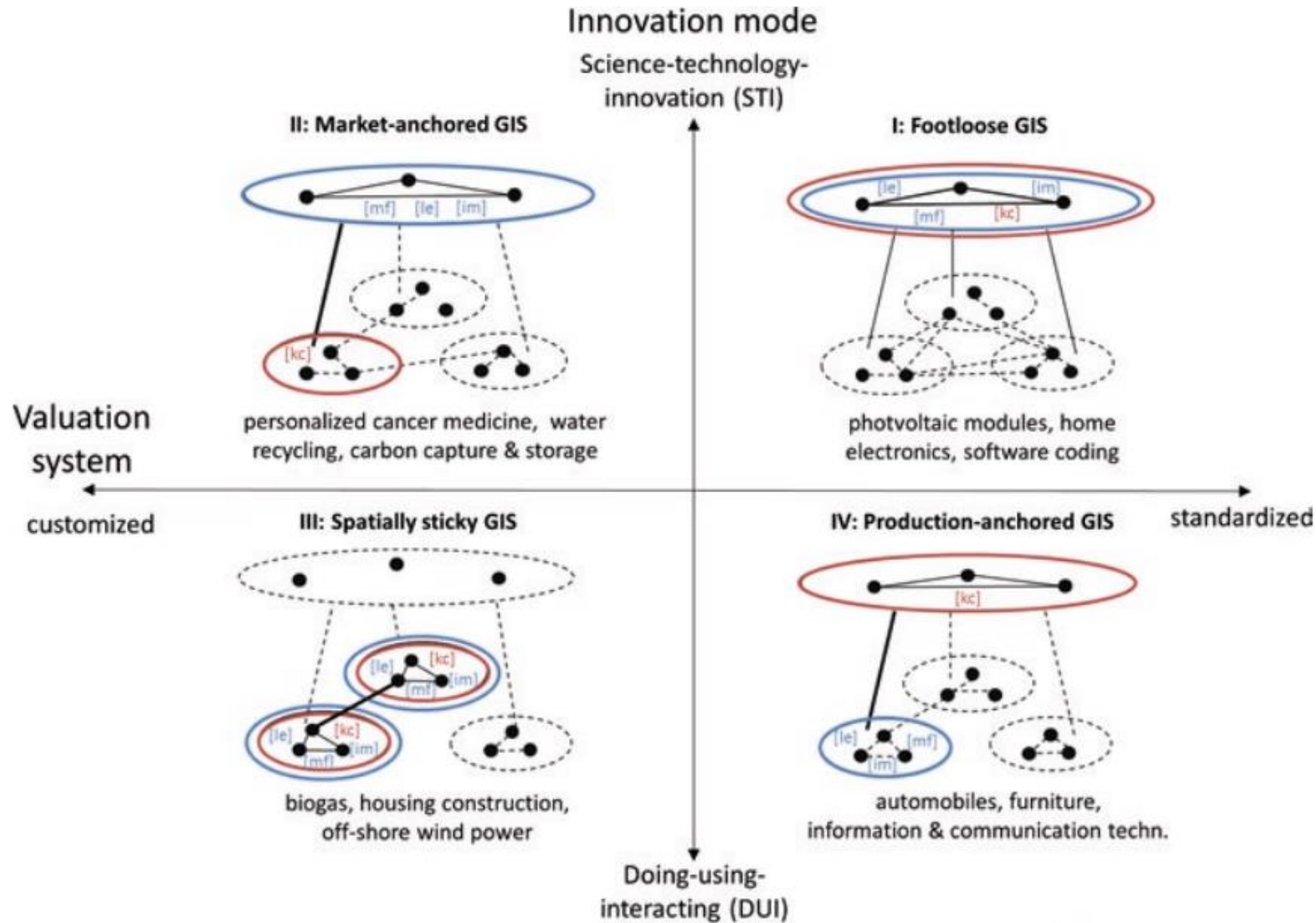
Indicator groups	Examples
Input	Research proposals, project fundings
Troughput	Patents, scientific publications
Output	Newspapers, social media, user forums

# Prioritizing along systems in the transition process

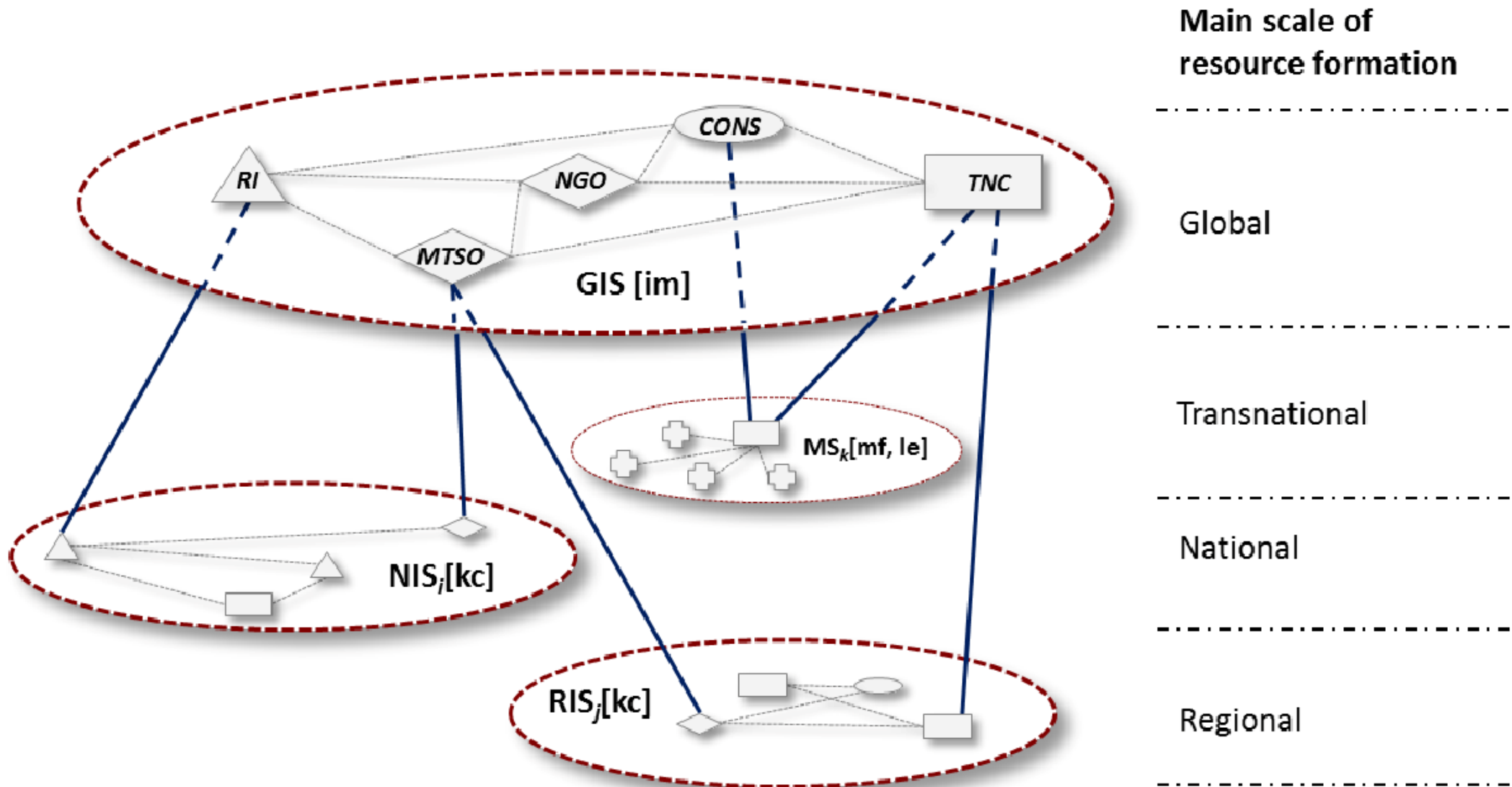




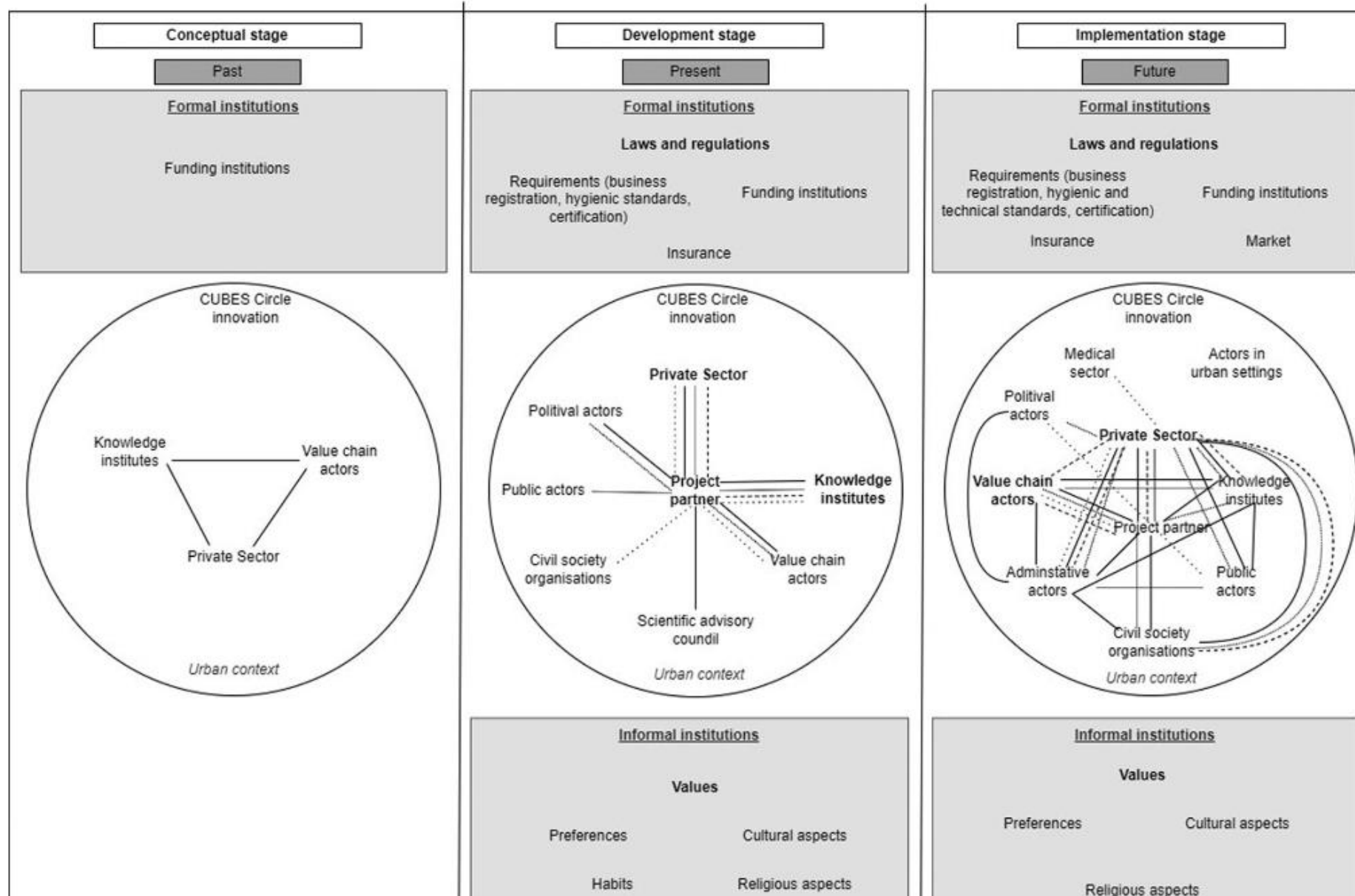
# Prioritizing regional, national, supranational factors



# Prioritizing regional, national, supranational factors



# Prioritizing factors along development phases

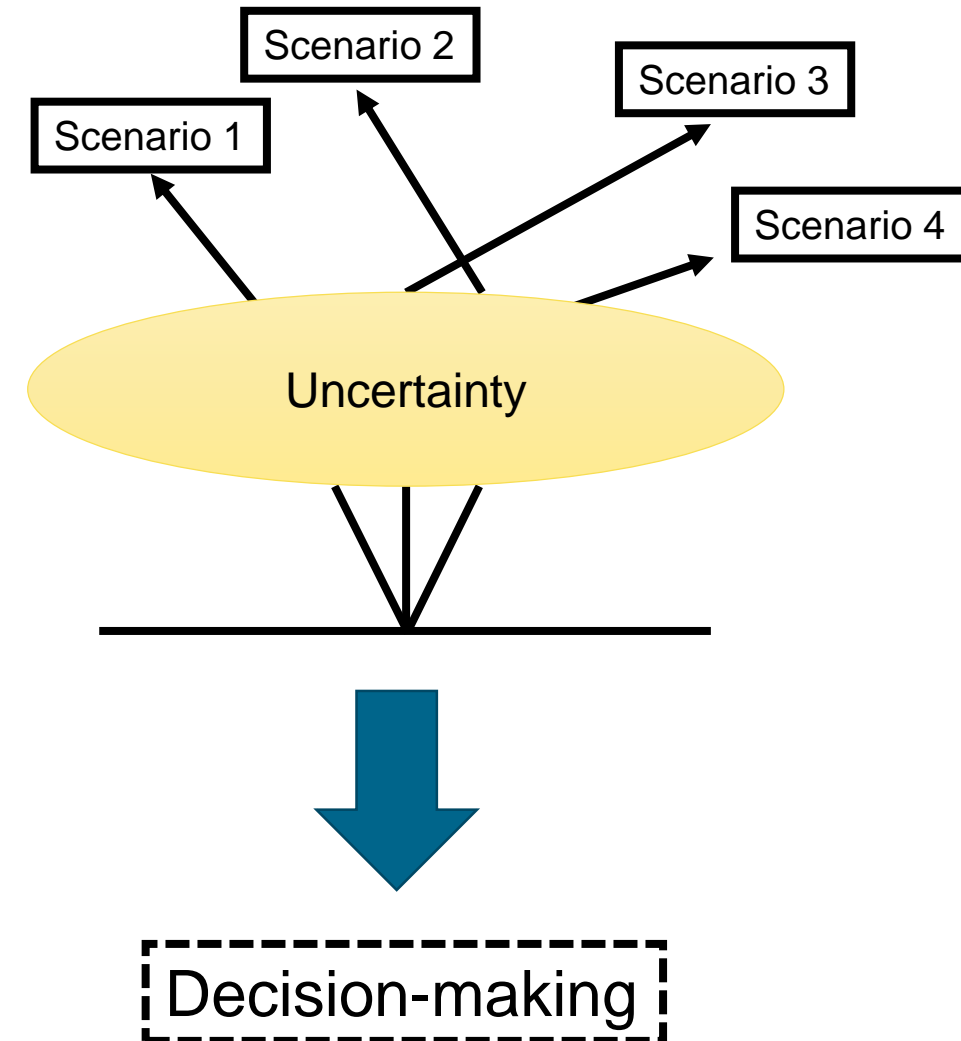




# SUMMARY AND OUTLOOK

# Summary and Outlook

- Text mining to close data gaps
- Framework for translation into decision-making needed (Müller 2022)
- **But:**
  - **Validation still difficult!**
  - **LLMs relatively slow and costly!**





# Outlook – LLMs?



## ➤ Advanced NLP-tasks

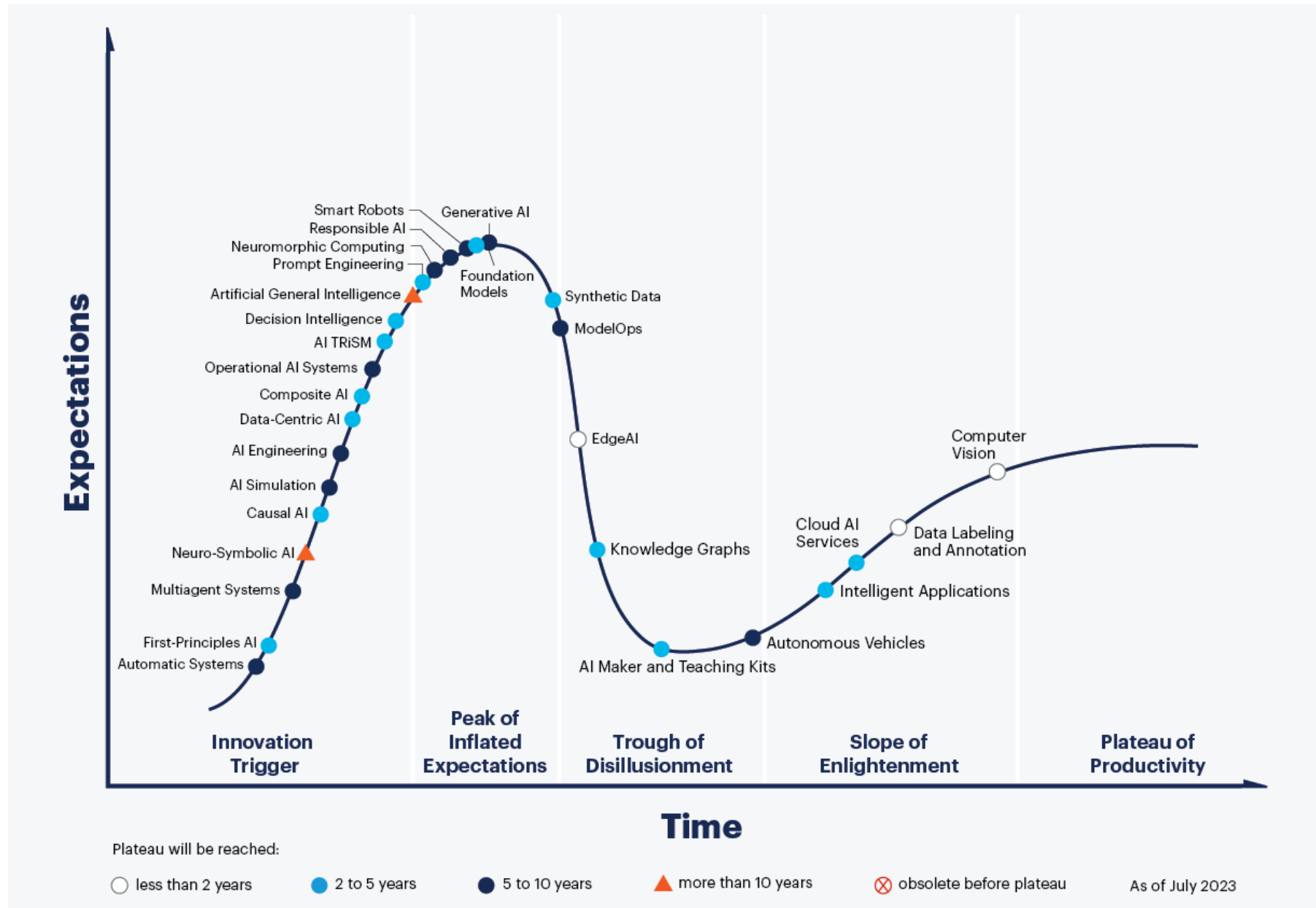
- Q&A
- Chatbots
- Text understanding?

## ➤ But trade-offs to traditional NLP-methods, e.g., blackboxing

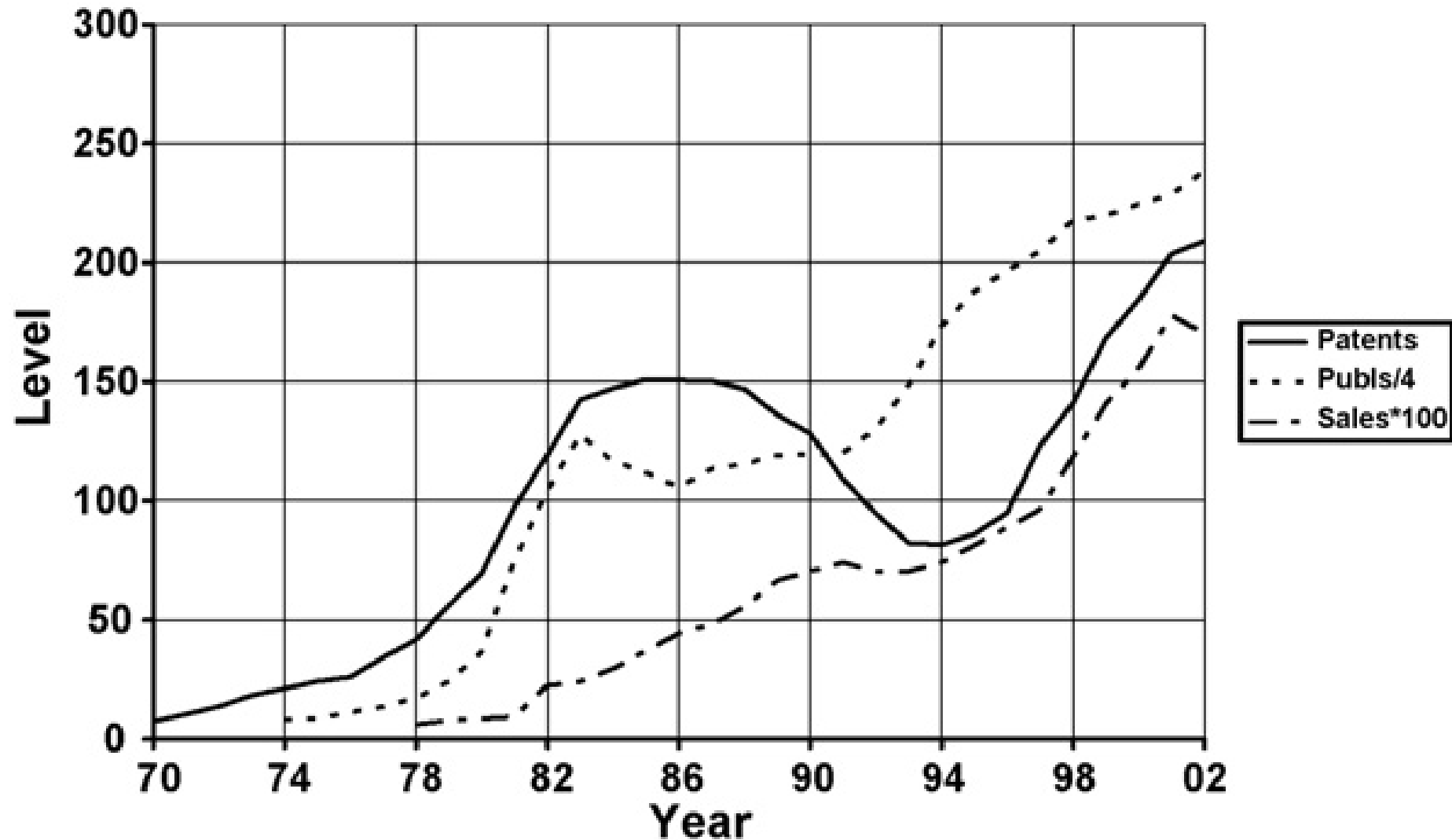
## ➤ **Substitute to innovation researchers?**



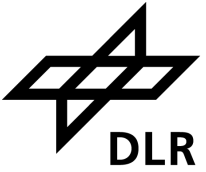
# Outlook – Gen AI (Gartner hype cycle 2023)



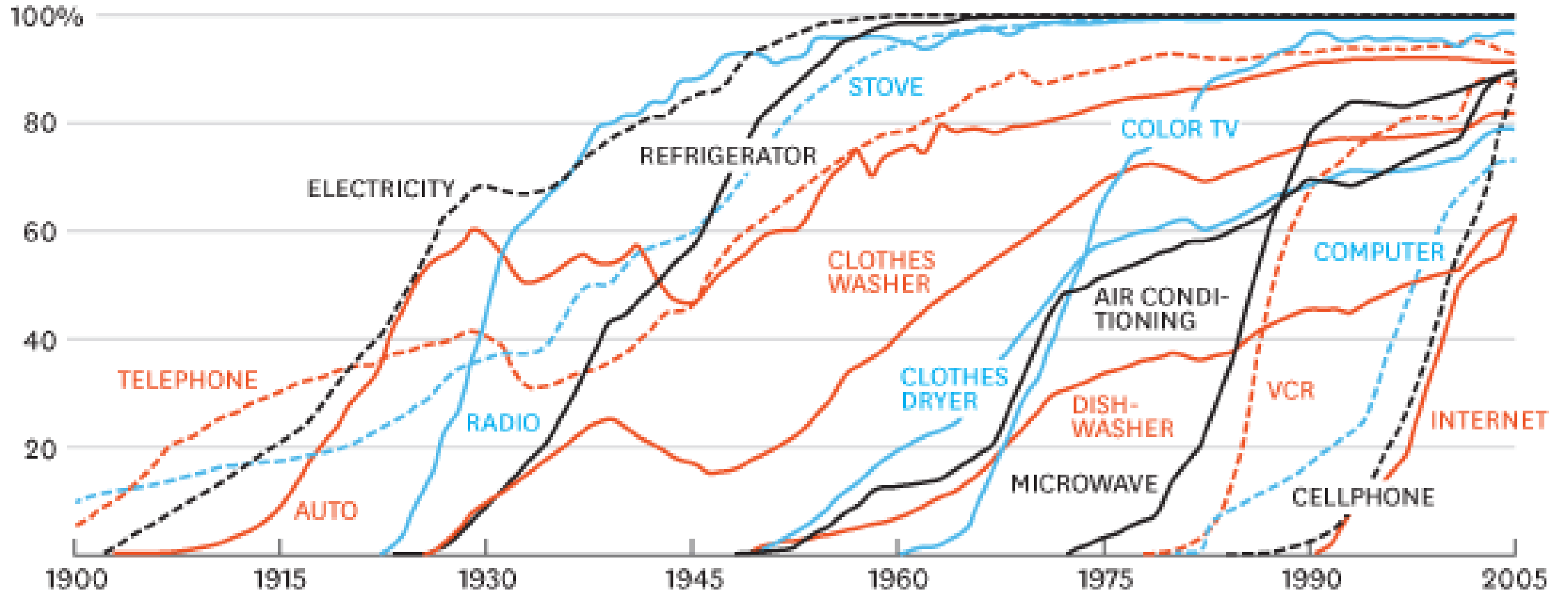
# Double boom-cycle in Robotics



# Heterogenous technology diffusion curves

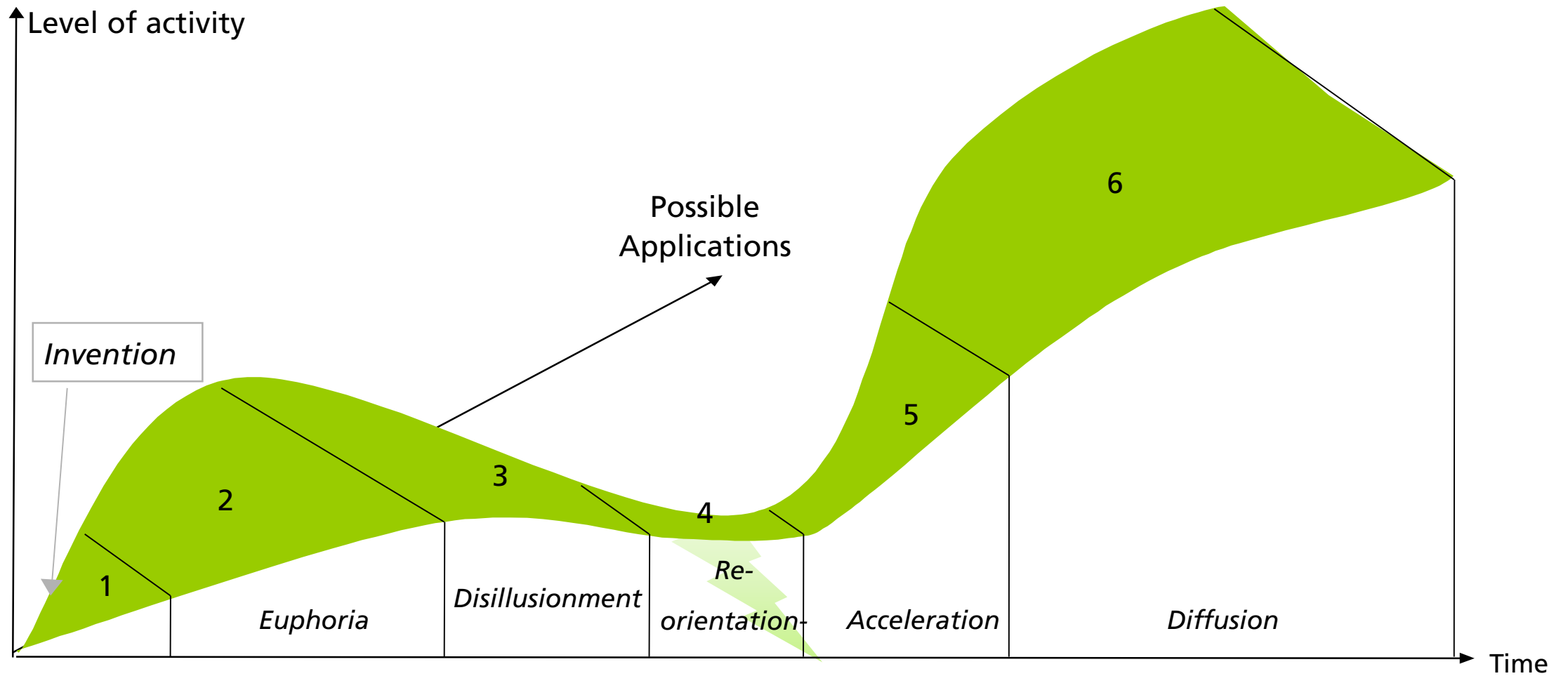


PERCENT OF U.S. HOUSEHOLDS



SOURCE MICHAEL FELTON, THE NEW YORK TIMES

HBR.ORG





# Selected References



- Bellstam, G., Bhagat, S. and Cookson, J. A. (2021) 'A text-based analysis of corporate innovation', *Management Science*, vol. 67, no. 7, pp. 4004–4031.
- Fagerberg, J. (2002, April). A layman's guide to evolutionary economics. In Industrial R&D and Innovation Policy Learning conference, Leangkollen, Asker (pp. 18-19).
- Hekkert, M. P., Suurs, R. A.A., Negro, S. O., Kuhlmann, S. and Smits, R. E.H.M. (2007) 'Functions of innovation systems: A new approach for analysing technological change', *Technological Forecasting and Social Change*, vol. 74, no. 4, pp. 413–432.
- Kim, J., Park, Y., & Lee, Y. (2016). A visual scanning of potential disruptive signals for technology roadmapping: investigating keyword cluster, intensity, and relationship in futuristic data. *Technology Analysis & Strategic Management*, 28(10), 1225-1246.
- Markard, J. (2020) 'The life cycle of technological innovation systems', *Technological Forecasting and Social Change*, vol. 153, p. 119407.
- Waldherr, A. (2012) 'The Mass Media as Actors in Innovation Systems', in Bauer, J., Lang, A. and Schneider, V. (eds) *Innovation Policy and Governance in High-Tech Industries: The Complexity of Coordination*, Berlin, Heidelberg, Germany, Springer, pp. 77–100.
- Weiss, D. and Nemecek, F. (2021) 'A text-based monitoring tool for the legitimacy and guidance of technological innovation systems', *Technology in Society*, 66, p. 101686. Available at: <https://doi.org/10.1016/j.techsoc.2021.101686>.
- Weiss, D. and Nemecek, F. (2022) 'A Media-based Innovation Indicator: Examining declining Technological Innovation Systems', *Environmental Innovation and Societal Transitions*, 43, pp. 289–319. Available at: <https://doi.org/10.1016/j.eist.2022.04.001>.
- Weckowska, D., Weiss, D., Fiala, V., Nemecek, F., Voss, F., & Dreher, C. (2024). Creating legitimacy for cultured meat in Germany: The role of social cohesion. *Environmental Innovation and Societal Transitions*, 52, 100871.

**Dr. rer. pol. Daniel Weiss**

*Research group leader „Market analysis, -  
design, and innovation strategies*

*DLR Institute for Transport Research (VF)*

***Daniel.weiss@dlr.de***  
***www.dlr.de/vf***

**THANK YOU!**