



Modeling AI Trust for 2050: perspectives from media and info-communication experts

Katalin Feher¹ · Lilla Vicsek² · Mark Deuze³

Received: 6 October 2023 / Accepted: 20 November 2023
© The Author(s) 2024

Abstract

The study explores the future of AI-driven media and info-communication as envisioned by experts from all world regions, defining relevant terminology and expectations for 2050. Participants engaged in a 4-week series of surveys, questioning their definitions and projections about AI for the field of media and communication. Their expectations predict universal access to democratically available, automated, personalized and unbiased information determined by trusted narratives, recolonization of information technology and the demystification of the media process. These experts, as technology ambassadors, advocate AI-to-AI solutions to mitigate technology-driven misuse and misinformation. The optimistic scenarios shift responsibility to future generations, relying on AI-driven solutions and finding inspiration in nature. Their present-based forecasts could be construed as being indicative of professional near-sightedness and cognitive dissonance. Visualizing our findings into a Glasses Model of AI Trust, the study contributes to key debates regarding AI policy, developmental trajectories, and academic research in media and info-communication fields.

Keywords Artificial intelligence · Info-communication · Media · AI Trust · Beliefs · Uncertainties · Generative AI

1 Introduction

Information Communication and Media (ICM) is a rapidly growing field that integrates society, culture, and technology (Fuchs 2009). Amid the hype surrounding artificial intelligence (Newlands 2021), professionals in ICM fields such as telecommunications, journalism, entertainment, marketing, social media and information technology now face the question of whether they can trust the technologies they help to create. Although human–robot communication is a critical topic in academic research over the last decade (Feher

and Katona 2021), research on the challenges and future visions of ICM systems are sporadic or under-represented in academic publications and strategic AI documents. This is why it is particularly relevant to investigate the visions of ICM professionals.

The key research question is how ICM fields are changing due to AI technology and what the key consequences are expected to be. Our study design was deliberately open to obtain a wide bandwidth of answers to the questions asking about the possible futures of AI. Sub-questions addressed the advantages, benefits, disadvantages and uncertainties of AI, and probed for reflections on all the changes this may bring. We also wanted to understand the transformation in the ICM process and related changes in human–social values vis-à-vis the emerging AI environment (Feher 2020).

This research is especially timely if we consider the rapid spread of AI-driven ICM phenomena such as conversational media, deepfakes, recognition systems, AI-driven audiovisual media, bot journalism, generative music, social media, recommendation systems or synthetic media (Hight 2022; Trattner et al. 2022; Hartmann and Giles 2020). Text and audiovisual content produced by AI can be easily accessible and persuasive, even if it may be biased, offensive or misleading (Illia et al. 2023; Jackson and Latham 2022),

✉ Katalin Feher
feher.katalin@uni-nke.hu

Lilla Vicsek
lilla.vicsek@uni-corvinus.hu

Mark Deuze
M.J.P.Deuze@uva.nl

¹ University of Public Service, Ludovika Ter 2.,
Budapest 1083, Hungary

² Corvinus University of Budapest, Fovam Ter 8.,
Budapest 1093, Hungary

³ University of Amsterdam (UvA), Turfdragsterpad9,
1012 XT Amsterdam, The Netherlands

implying trust issues as well as related concerns over authorship and verifiability. The outputs of current AI services are between 50 and 70% accurate, while they also produce false or invented things, so-called “hallucinations” (Lin et al. 2021). Thus, authenticity, deception and trust have become fundamental issues (Glikson and Asscher 2023; Hancock et al. 2020), also in policy research (Pierson et al. 2023).

AI-generated ICM and generative AI increasingly reach a broad audience globally (Pavlik 2023; Kemp 2022) and influence decision-making, investment and policy development, understanding the visions and expectations of key professionals involved in this shift is crucial. If we can assume that experts trust the technology enough to invest their resources into the future of AI, even if they express serious concerns (Feher and Veres 2022), they can be seen as “technological trust mediators” (Bodo 2021), prominently involved in the socio-cultural construction of AI (Feher and Katona 2021). Thus, the study assumes that ICM experts are ambassadors of AI technology in terms of awareness of social and ethical issues, shaping the field for years to come.

According to the analysis of the survey, the participants’ visions can be mapped as a Glasses Model of AI Trust, representing their way of balancing hopeful beliefs, growing concerns, and overall uncertainties. This model visualizes how key values related to the potential of personalized and unbiased ICM operate in relation to human values, observable limits of AI, and structural reliability issues. The dynamics of believed and uncertain effects as represented in our model suggest that the future vision as expressed by our study participants is optimistic, where problems tend to be interpreted as opportunities and even as advantages. In this way, the Glasses Model of AI Trust allows us to understand a dominant operational logic informing current and upcoming innovations.

2 Theoretical framework

In the background of all-encompassing info-communication systems (Kovtun et al. 2022), mediatization (Hepp 2020; Feher 2022), platformization (Van Dijck 2021) and artificial intelligence technologies (AI) develop and spread to extend human-machine communication (HMC, Guzman et al. 2023), AI-mediated communication (AI-MC, Hancock et al. 2020), generative AI (Pavlik 2023) algorithmic and proxy media (Blanchett et al. 2022) or super-human intelligence for interactivity (Guzman et al. 2023). These trends engender diverse computational and AI agent roles encompassing senders, receivers, communicators, mediators, producers, editors, authors, creators, designers, analysts, distributors, moderators, and fact-checkers. Although it is useful to delimit the research focus (Hancock et al. 2020), expansive research allows the exploration of a vision of a distant future

for free, abstract, or specific associations. Especially in the case of complex and rapidly evolving AI technology with influenced trust attitudes and decision-making, among others in the case of media consumption (Araujo et al. 2020).

Although AI is a set of sophisticated agents, it tends to be interpreted as a black box technology: the machine operation remains largely hidden from human comprehension (Rassa-meeroj and Wu 2019; von Eschenbach 2021). The concept of black box technology is still dominant and creates several uncertainties—even if a concept of glass box transparency is emerging, promising a more understandable machine behavior for building trusted AI (Toy 2023). At the same time, AI is a driver of transforming digital services for numerous areas of applications. Humanity and AI technologies mediate one another via interaction and collaboration (Borsci et al. 2023; Verbeek 2015), despite such uncertainties and critical concerns. The datafication and deep mediatization of society are rewriting the political economy of society and IT industries (Brevini and Pasquale 2020), as well as everyday life, leading to extremely high expectations or future nightmares expressed by different stakeholders (Mansell and Steinmuel-ler 2020). In doing so, the responses and expectations surrounding AI mirror earlier technological transformations, which inspire both utopian and dystopian scenarios (Feher and Veres 2022).

Of seminal concern here is what drives AI regarding ICM with the branches of several technologies to accelerate existing processes (Hui 2021) and introduce new pathways. We can distinguish at least two driver functions in the case of ICM. First, AI communicates, interacts and audio-visualizes (Fletcher 2018), thus it makes itself perceptible to humans. In short, it represents itself via ICA as a surface of the technology. This process is an elementary way for humans to sense and reflect on the technology they create directly. Thus, the products of AI-driven ICM are one of the most controversial fields. Second, image, sound, and text have become AI-mediated in datafication and algorithmic operations (Ellis and Tucker 2020). The technology does not only support information flow in this way but also restructures the concepts of previously known new media and computer-mediated communication (Guzman and Lewis 2020). Big/smart/synthetic data, machine and deep learning, neural and recommendation networks train algorithms to determine what is popular, personalized or fake in the temporary relevant systems.

Considering these two driver functions, we can theorize the meaning of AI in the particular context of ICM fields following the approach to media as socio-technical systems, whereby media are seen as the “intersections of technical knowledge, humanistic investments, social relations, economic models, political stakes, and aesthetic expression that people use to understand and shape their lives” (see: uscmasts.org). The term “info-communication” also

deliberately comprises both computer-based and human-driven processes in interaction (Targowski 2019). Therefore, these mutually foundational and complementary processes both enable and constrain human civilization and social inclusion primarily via advanced ICM technologies. Consequently, the application fields of ICM processes can be found throughout various sectors of industry and society.

AI-driven ICM flourish in synthetic worlds (Gunkel 2019) where previous media and communication operations seem less sustainable (Chan-Olmsted 2019), and where the identification of (degrees of) reality significantly depends on machine learning (Waisbord 2018). Cost-effective and productive operation is an industry-wide expectation from the buoyancy force of AI (Mustak et al. 2023; Preu et al. 2022; Wirtz 2020), even if trustworthiness, reliability, and bias are all at stake. In parallel, a key requirement is to control (and clean up) datasets, fight disinformation and foster truly diverse, inclusive and reliable content (Georgieva et al. 2022). These transformations support benefits and trigger uncertainties, especially as the above-mentioned driver functions reveal challenges of socio-technical issues with inequalities (Holton and Boyd 2021), system biases (Rawat and Vadivu 2022), cultural-economic colonization (Bell 2018), data colonialism (Couldry, and Mejias 2019), data-driven surveillance with privacy issues (Fossa 2023) and political destabilization through fake campaigns (Borsci et al. 2023). The interactivity and virality of personal assistants, virtual influencers, AI-produced content and art, or deepfakes leads to fundamental questions such as how information sources can be evaluated or how they will add value to an existing ICM process or service. Furthermore, the livelihood of many practitioners in creative roles in these fields seems to be at stake (as exemplified in the 2023 writers' strike in Hollywood).

Accordingly, ICM experts working in various areas face several challenges as the complexity of their field of expertise grows (Swiatek et al. 2022). Besides being active in their fields, they also contribute to setting directions for the future – their own as well as that of generations to come. Given the fact that their responsibility is inherently connected to minimizing human errors and maintaining trust in AI systems, the pressure is on (Borsci et al. 2023; Ryan 2020; Amaral et al. 2020). This study aims to understand their current interpretation, future projections and the dynamics of their visions.

3 Assumptions and the research questions

Considering the theoretical framework and its key components, four assumptions frame our project:

- AI fundamentally transforms ICM systems in several ways (Ellis and Tucker 2020; Guzman and Lewis 2020; Chan-Olmsted 2019; Gunkel 2019; Fletcher 2018; Waisbord 2018).
- Multifarious benefits are available, from cost-effective operation to productive work processes (Mustak et al. 2023; Georgieva et al. 2022; Preu et al. 2022; Wirtz 2020).
- Various uncertainties and dangers exist, from fake media and systemic bias to the possibility of abuse (Feher and Veres 2022; Borsci et al. 2023; Rawat and Vadivu 2022; Holton and Boyd 2021; Bell 2018).
- Socio-cultural values are discussed in parallel to trust issues (Borsci et al. 2023; Feher and Katona 2021; Ryan 2020; Amaral et al. 2020).

The assumptions were transformed into a series of explorative research questions allowing for free associations.

RQ1. How are ICM processes changing due to AI technology?

RQ2. What are the benefits of the change?

RQ3. What disadvantages shape the changes?

RQ4. How do socio-cultural approaches change, and how does this process affect trust in AI technology?

4 Method and sampling process

We conducted the survey online with a sampling approach based on sufficient diversity of member characteristics and the four research questions guiding our project (Jansen 2010). In addition to a first set of demographic questions, participants received two or three complex questions per week and were given four weeks in which to answer them. This schedule allowed the participants to (1) plan their time (2) be engaged in the process (3) have time to recall their knowledge or check academic/professional sources for more detail (4) give complex answers to broad questions on a weekly basis.

With the goal of obtaining a diverse expert sample in a multidisciplinary field, the opening round of survey questions addressed world regions, gender, profession, sector, years of experience, and background (academia, business, policy-making or NGOs). These are all relevant to understanding different perspectives, especially in the case of a relatively small sample. A second round of questions was derived from the RQs, broken down into detailed subquestions while accounting for the two time dimensions of present and future:

1. What is the meaning of AI technology to you? (Question of AI terminology)

2. How would you imagine your life in 2050 when you communicate, entertain, and acquire information via AI-generated or AI-supported services? (RQ1)
3. Predominantly, how does AI technology shape the media, information and/or communication process inside and outside of your industry? What will change in this field in the future, till about 2050? (RQ1)
4. What are the key benefits of AI applications in the case of media, information and/or communication technology? What will change in this field in the future, till about 2050? (RQ2)
5. How does AI technology support media, information and/or communication production or consumption? What will change in this field in the future, till about 2050? (RQ2)
6. What are the disadvantages of AI applications in the case of media, information and/or communication technology? What will change in this field in the future, till about 2050? (RQ3)
7. How does AI technology generate issues and uncertainties in the future from media, information and communication technology? What will change in this field in the future, till about 2050? (RQ3)
8. Does AI technology change our social-cultural values and norms via the media and information-communication technology? If yes, how? If no, why not? What will change in this field in the future, till about 2050? (RQ4)
9. Does AI technology change our social-cultural values and norms via the media and information-communication technology? If yes, how? If no, why not? What will change in this field in the future, till about 2050? (RQ4)

More than 300 experts were selected and invited based on their LinkedIn profiles and 42 participants joined from all world regions. Of these, 25 participants engaged with detailed answers throughout the sampling process for four weeks. Every week, they received a link with survey questions, which they had one week to answer via a Google Form. The participants joined voluntarily and anonymously, consenting to a GDPR-compatible contribution to the sampling which was conducted in 2022.

The respondents summarized their experience, observations, practice, visions, and examples. We rewarded the 25 participants for their engagement with an executive summary of the results. This ratio is suitable because of the complex topic, the duration of the survey (several weeks) and the lack of additional rewards.

After compiling the whole dataset in one database, data cleaning was applied to eliminate mistyping and other language errors. The final dataset contains 21,441 words in the whole sample and 18,831 words with just the answers of the 25 participants. NVivo qualitative software (version 11) was applied for detailed analysis. Horizontally, the answers

were automatically coded in terms of the survey questions, while manual codes were applied vertically to identify and synthesize the most common topics, considerations, beliefs and uncertainties. Two authors independently performed manual coding line-by-line for the credibility of the analysis. The cross-checked coding also supported making memos about the most relevant and agreed-upon patterns in the texts for interpretative analysis. We detail the survey results and analysis in the next section.

5 A diverse sample

The sampling method yielded diverse member characteristics (Jansen 2010). The 42 participants of the first week (hereafter entire cohort) came from all world regions and had an average of 5–10 years of work experience in several fields (see Table 1). From the entire cohort, 25 participants as the subset cohort engaged for the whole sampling period, representing primarily female and academic respondents, all age groups and almost all world regions from about half of the disciplines and sectors (see Table 1). Despite the diversity of the sample, there were surprisingly homogenous answers with similar visions. Only the African respondents highlighted quite different perspectives, raising some significant concerns about the uneven impact of AI. Since there were more female respondents gendered concerns were also available with examples of a high rate of male AI professors, under-represented female data, and the negative impact of AI-generated images in porn and social media. In addition, experts from academia and NGOs were the most sensitive to issues of ethics and trust.

6 Findings

6.1 Terminology and future vision of AI

Respondents from the Entire Cohort defined AI as either advanced technologies or automated systems – or merely their fetishization. In both cases, machines are trained to achieve goals like problem-solving and predictive analytics. A few emphasized instead the machine's capacity for learning and mental tasks.

Although not all participants are likely to be alive in 2050, they formulated a vision of a supportive AI system to work instead of humans as well as for humans. Advanced search engines and wireless connections are expected to enable direct access to all information. Sociable AI combining virtual assistants and humanoid robots, known as Computers as Social Actors (CASA), will enhance interactivity. Everyday routines will be assisted resulting in AI-generated personalized experiences significantly affecting

Table 1 The sample

Cohort	Gender	Age	Region	Background	Experience	Sector/discipline
Entire cohort:	Male 18	18–29: 7	Europe 20	Academia 23	<5 years 12	AI Ethics
42 experts	Female 24	30–39: 16	North America 9	Industry 16	5–10 years 15	AI Policy and Governance
		40–49: 12	Africa and Middle East 5	NGO 2	10 < years 15	AI Research
		50–59: 4	Asia 4	Policy making 1		Arts, humanities, philosophy
		60+ : 3	Asia–Pacific 3			Behavioral Economics
			Central and South America 1			Biotechnology
					Computer Science/engineering	
					Communication, marketing, media	
					<i>Futures and Foresight</i>	
					<i>Fintech and investment</i>	
					<i>Higher education</i>	
					<i>Innovation/tech consulting</i>	
					Knowledge/IT Management	
					Law	
					Machine Learning and data analytics	
					<i>Project management</i>	
					<i>Public health</i>	
					Speech and Language Processing, NLP	
					Tech-entrepreneurship	
					TESL	
					UX	
Subset cohort:	Male 10	18–29: 2	Europe 12	Academia 16	<5 years 7	AI Ethics
25 experts	Female 15	30–39: 10	North America 4	Industry 7	5–10 years 7	AI Policy and Governance
		40–49: 6	Africa and Middle East 3	NGO 1	10 < years 11	AI Research
		50–59: 4	Asia 3	Policy making 1		Arts, humanities, philosophy
		60+ : 3	Asia–Pacific 3			Behavioral Economics
						Biotechnology
					Computer Science/engineering	
					Communication, marketing, media	
					Knowledge/IT Management	
					Machine Learning and data analytics	
					Speech and Language Processing, NLP	
					Tech-entrepreneurship	
					TESL	

human decision-making and emotions. Moreover, a few respondents envisioned perfectly personalized AI bubbles catering to individual human needs and desires.

Well-being, security, and sustainability via technological developments were the focus of the answers. Most respondents discussed technological determinism to avoid dystopic sci-fi scenarios and political misuse. This approach stresses the need for AI governance and ethics to apply human–AI collaborations. Beyond the advantages and general changes, some concerns raised were dependence on machines, risks of an outage, information wars, growing inequality, limits in accuracy, reduced human communication, growth of AI-like behavior in entertainment, security issues or AI-produced carbon footprint. The majority were rather optimistic and believed in the power of NGOs, art, youth empowerment and the rise of fringe regions with AI-driven decolonization and recolonization.

In conclusion, experts from the entire cohort offered a broad definition of AI affecting all aspects of social-political

life, with AI potentially supporting human prosperity significantly, although they also raised a few concerns. Strong words, such as “hope” and “believe”, support the positive approach. Their technology-based vision of the future is fundamentally deterministic and rather optimistic.

6.2 Horizontal results: answering RQs

The subset cohort formulated a present-based vision of the future. Accordingly, two parallel worlds were defined for RQ1 by the experts:

1. automated information and data processing for algorithmic news and fact-checking;
2. misinformation, cheap fakes, semi-true news, and the misuse of AI technology.

They listed several AI-powered ICM methods and tools, including biometrics, social media, and data analysis, with

the most popular applications, such as Netflix, Twitter, Facebook, TikTok or the deepfake video of Volodymyr Zelensky. One of its respondents concluded that “AI is both a medium and an entity/actor”, which connects with literature such as Ellis and Tucker (2020) and Guzman and Lewis (2020), confirming the relevance of the research.

All participants agreed that there is a key conflict between democratic values and social control, although they could not predict how this dynamic will change. Yet, the majority of the respondents had a positive future vision of universal access to trustworthy AI in a democratic way, supported by AI-driven control and AI-generated mitigation of potential adverse effects.

A few participants proposed the idea that digital communication with nature will also be a source of AI improvements if interactions between animals and plants are connected to big data and smart systems. Thus, we formulated the term “nature listening” for this feature, based on the practice of “social listening” (Stewart and Arnold 2018) in terms of the advanced monitoring of social media activities by platform companies. The answers to RQ1 are summarized in Table 2, Row 2.

Answering RQ2 about potential benefits, respondents’ “imagination” centered on all-the-time available information from cost-effective production to preventing negative impacts. At present, touchpoints support the funneling of experiences by chatbots or virtual assistants, such as Amazon Alexa and Apple Siri. In addition, virtual media and art creators impact human perceptions from music to visual images, such as AIVA, Dall-E, ChatGPT, Midjourney or PearAI.Art. The key challenges are legal, ethical and reliability issues, media fragmentation, social inequalities, the reproduction of colonialism or decreased creativity. Recalling the literature review, almost all key issues are repeated here.

The subset cohort participants felt that, by 2050, AI-driven ICM will influence how users experience their all-surrounding circumstances beneficially if human narratives are accelerated and augmented. They highlighted an omnipresent AI-media with demystification of the ICM process if every human agent is wired. Open AI options were mentioned most often in connection with reaching a healthy democratic future with digital sovereignty and the power of communities. See more details in Table 2, Row 3.

RQ3 focused on the disadvantages and negative aspects of AI-driven ICM (Table 2, Row 4). According to the respondents, AI-driven ICM has drawbacks, primarily related to uncertainties caused by unreliable and biased data. Misinformation, filter bubbles, non-filtered information overload and machine-dominated communication and news production exacerbate these uncertainties. Without proper regulations and ethics, this can decrease industry

value and harm societal processes. Accordingly, two scenarios were predicted most frequently for 2050:

1. Lack of diversity and equity in the AI-driven ICM field, resulting in semi-fake news and the suppression of freedom of speech due to redefined censorship, centralization of ownership and exclusive gatekeeping functions
2. More reliable and accountable AI systems through transparency, regulation and the right to be forgotten. Thus, legal frames and innovation should reduce uncertainties, even if one female academic participant in AI Marketing from Europe emphasized: “*Only one thing is certain: uncertainty. But what’s also assured is that organizations will unlock value from AI and encounter challenges in unexpected places.*”

RQ4 asked about trust in AI technology and socio-cultural standards in the ICM process. Respondents noted that while we are only beginning to understand the impact of AI on society, economy, and culture, fundamental issues such as profit maximization and societal asymmetry remain unchanged. AI cannot independently alter socio-cultural values and norms but might amplify or repress them, leading to new forms of abuse and violence. A few participants imagined socio-cultural changes but without specifications, such as.

“Meanings are being affected through our experiences with AI and our interactions with social chatbots and other sociable AI. These meanings become part of our social discourse and socio-cultural values, norms and codes” — A male academic participant in communication, North America

Others mentioned two more examples of social change:

1. Religions may become more open and tolerant in communicating on AI-driven platforms, inspiring different practices worldwide.
2. Data sets can act as colonizing languages and cultures, thereby decreasing diversity, exacerbating existing inequalities, and social biases and resulting in further fragmentation.

Participants envisioned that by 2050 that bots, tokens, and synthetic accounts will drive human perception and communication resulting in a more pervasive role of technology throughout all aspects of everyday life. As one respondent expressively formulated:

“It may sound dystopic, but we are not very far from an internet overflowing with automatically generated content, in which the opinions of real people will be buried under a mountain of generated tokens. Another

Table 2 Summary of the answer to RQs

RQs	Current status	ICM in 2050
RQ1 How are ICM processes changing due to AI technology?	AI-driven or AI-abusing ICM Democratic values are in conflict with control and manipulation Freedom of speech or manipulation are under debate	Universal access, democratically available information and automated personalization Operation efficiency, enhanced decision-making, mistake prevention and decreased misinformation Trust in AI through human values, community platforms, multi-tasking and “nature listening”
RQ2 What benefits facilitate the change?	Automation, data analysis, personalization, cost-effective production, and recommendation systems AI-filtered/-controlled misinformation and fake media Touchpoints to experience the system, virtual media/art creators	Captology-based insight, data trust, unbiased systems Demystification of ICM, digital sovereignty, personalized enclaves Omnipresent ICM systems with virtual influencers
RQ3 What disadvantages shape the changes?	Biased systems, non-reliable data, lack of transparency Unclear sources of information, fake media, infoxication, echo chambers, online scams, filter and consumption bubbles, media fragmentation Social inequality, an illusion of connectivity	Transparency, accountability, regulation for verified data and non-biased services Semi-true news as a norm, behavioral data systems in innovation, incognito mode, erasable information, deepfakes Reduced human contact, suppression of freedom of speech, changing human thinking
RQ4 How do socio-cultural standards change, and how does this process affect trust in AI technology?	Options for adaptations but anchors of colonization in data sets Behavioral targeting and bubbles, fragment society and culture For a higher level of trust, machines should be trained for human morality	New generations build their own trust-as-practice, but democratic and anti-democratic services compete Unbiased systems are available but technology determinism is confirmed More research to support value-driven ethical and responsible AI services

way in which AI affects human opinions is through recommendations, these focus on eliciting certain feelings, such as anger, which make users engage more with the content in hand. This can have effects such as the growing polarization we have experienced in the last decade.” — A male non-academic participant in computer engineering, Europe

Extreme personalization and profiling are expected to decrease cause-and-effect-based human thinking and automate intuition in the ICM field, resulting in ethical issues. While some fear extreme control or misleading information, others do not believe in significant changes in society and culture. Trust in AI technology is not a major worry, as social learning about changes is ongoing and primarily determined by technology, and new generations are already familiar with AI environments.

6.3 Vertical results: beliefs, uncertainties, trust

The most common approaches and considerations of the subset cohort became available for detailed interpretation by manual codes. According to the cumulative results, only a few experts are truly skeptical, expressing their technopessimism and surprisingly, only one example addresses reliability issues:

“I will not believe any algorithm and do not consider it a reliable source. The press will not be a reliable source of information, nor will companies [offer] sound opinions, nor will the reports based on data collection.” — A female non-academic participant in entrepreneurship, Africa and the Middle East

The majority believe in what they are doing now for the proper use of AI and look to the future confidently. Even if they articulate uncertainties, they believe AI can support numerous ICM processes effectively — from entertainment to the news. Illustrating the most fundamental driving force of techno-optimism:

“AI will simplify large amounts of data to crisp information adaptable to a form of media, person and place. A place where machine and human can work hand in hand.”

— A female non-academic participant in PR and corporate communication, Asia–Pacific

The other driving force is a negative or skeptical approach, informed by plenty of uncertainties if the intention is the misuse of technology. Uncertainty primarily amplifies trust issues. However, beliefs and uncertainties can be two-fold arguments also highlighting the necessity of a reliable technology:

“Like all technologies, it creates uncertainties in how it can be used for harm instead of good in society both intended and unintended. For example, it can suppress freedom of speech very effectively online by deleting specific subjects or ideas. I believe that by 2050 state/company-sanctioned misuse will see more adoption so effective regulations will be put in place for the fair use of AI.” — A male non-academic participant in information technology, Europe.

OR

“I believe that one of the disadvantages of AI is portrayed as an advantage: the scale of AI applications. While scalability per se offers viable solutions for a wide range of applications, it also poses risks to diversity in algorithmic technologies, as well as to local and context-sensitive AI systems that are needed in fields like media, information and communication technologies.” — A female academic participant in media studies, Europe.

Even if the pros and cons are not always so directly next to each other, during the manual coding, it became clear that arguments were primarily driven by personal beliefs (and corresponding uncertainties). Other possible logics, such as advantages vs. disadvantages or utopia vs. dystopia were not reflected. People’s beliefs are the dominant forces for sensemaking and forecasting, resulting in generally neutral or outright positive future paradigms. Uncertainties were expressed less negatively or skeptically in this context. Instead, unknowns about the trajectory of AI seem to facilitate critical thinking, asking questions and rationalizing dominant beliefs and hopes.

A further noticeable result is that a significant proportion of the subset cohort participants argued that they could trust the technology because an AI-to-AI era is coming. To the participants this means that AI solutions will answer problems and threats generated by AI. Trust is built by AI developments for humans. An illustrative example is available below:

“I believe that only AI can help us in the fight of disinformation and we had a good example with the Covid-19 case on Twitter, Instagram and Facebook, for example: whenever the platforms recognized that content was related to Covid-19, it’d signal it to its users and this didn’t only refer to textual information, but AI was able to scrape through image content that was Covid-related and modulate its reach accordingly.” — A female academic participant in the behavioral economy, Europe

In many cases, a trusted future, as well as a definitive hope, is detectable in their replies. The word “hope” mainly appears in the statements of middle-aged European

participants with an industrial background and 5 to 10 years of experience in AI. Their hopes further strengthen personal beliefs, which tend to come with a highly positive and reassuring vision, such as:

“By 2050, I hope AI gives everyone the same opportunity to be great.”

— A female business participant in AI Ethics, North America

OR

“Hopefully, the technology of information source verification will develop to the point where these issues may be successfully tackled and resolved.”

— A male academic participant in philosophy and visual communication, Europe

Summarizing these results, a unified AI future was envisioned by almost all participants of the subset cohort as their beliefs and hopes for AI impact in the ICM process by 2050. This way of sensemaking and forecasting highlights a reciprocal dynamic between personal beliefs and expressed uncertainties. In addition, the high word frequency of hopes, beliefs, uncertainties and their synonyms suggested moving forward in the vertical analysis to analyze the contexts of these words manually. After cross-checking by the authors, beliefs and uncertainties could be separately coded and paired. By the context analysis, this result is directly connected to trust issues or trusted AI by hopes and critical thinking. These beliefs and uncertainties in a trusted AI context are critical when shaping policy agendas and developmental trajectories (Gross et al. 2019).

Accordingly, we propose the *Glasses Model of AI Trust* to summarize the essence of how this dynamic generates a future vision by the experts participating in our survey (Fig. 1).

The “glasses model” is a conceptual framework to understand how trust is balanced between beliefs underpinned by hopes and uncertainties fostered by critical thinking. Within this model, trust emerges as the fulcrum, maintaining a balance between these two contrasting elements resulting in dynamic strategies. Through the strategically combined lenses of beliefs and uncertainties, the experts are, on one hand, involved in shaping technology

and, on the other, see future alternatives. Their personal beliefs tend to strengthen trust in AI applications, while their uncertainties foster critical thinking without veering toward dystopianism. Trust in AI represents a balanced position or agreement that the technology can be ultimately reliable.

Our findings indicate that the key drivers of the model are hope and critical thinking. The participants hope AI integration will ultimately harmonize with human cultures, societies, and natural ecosystems. They believe that technology-driven effectiveness enables a greater focus on human creativity and imagination, while also fostering fairness, inclusivity, and the promotion of human rights. To illustrate this point:

“I envision a future in which the models upon which AI is developed will be perfected and – if regulated ethically and morally – will help us find a balance between freedom of speech, inclusivity, and the eradication of fake-news.”

A female academic participant in behavioral economics, Europe.

The other driver, critical thinking, is deployed when considering how AI might change job market dynamics and accelerate distrust. Two quotes that illustrate such concerns:

“Some jobs are going to be replaced. Some will disappear. And some new jobs will come along. Human workers need to enhance the core competitiveness.”

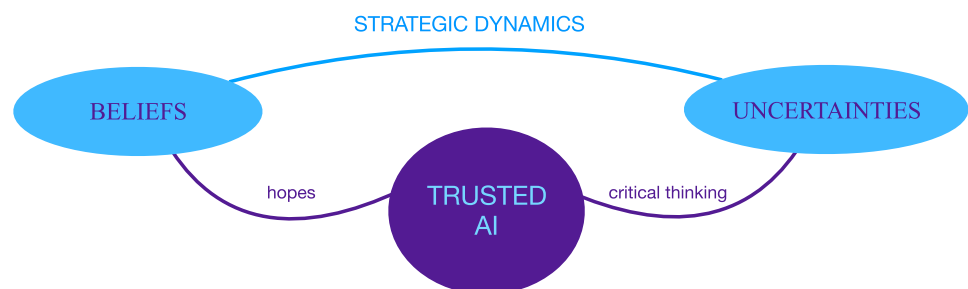
A female academic participant in communication and media, Asia

“There are concerning issues related to how deep-fakes/synthetic media can promote distrust and misinformation, especially in today's infodemic and post-truth world. Introducing even more information synthetically, such as AI is wont to do, could worsen these issues.”

A female academic participant in AI policy, Africa, and the Middle East

Hope and critical thinking can be seen as coping mechanisms in striking a balance between beliefs and uncertainties, specifically for appreciating the role of trust in AI.

Fig. 1 Glasses Model of AI Trust



Belief-based visions focus on the expected democratization of global conversations as language barriers disappear:

“By 2050, I believe people will break language barriers [with the help of AI]. Language translation will probably be solved by that time, with services automatically translating the language accurately in real time. The internet will be more integrated as the language barriers fade away.”

A male academic participant in NLP and IT, Europe

On the other hand, for example, uncertainties take into account the challenges of polarization, disinformation, and anti-democratic movements amplified through AI applications:

“AI systems could be utilized for anti-democratic purposes such as citizen surveillance and intimidation of activists and dissidents. We need cross-cultural AI ethics that take these dangers seriously.”

A female academic participant in marketing, Europe

Taken together, the beliefs and expressed uncertainties of AI experts give rise to strategic dynamics regarding (the future of) AI, balancing optimism with realism and ambitious yet practical guiding AI developments. These dynamics integrate beliefs and foster potentially innovative solutions with regard to uncertainties about societal and environmental impacts. As our study participants reported, all of this may enhance trust through transparent, realistic expectations of AI’s capabilities and limitations. One of the participants stated:

“An ethical foundation for tech is a must for humanity. Cross-machine learning will empower humans even more, subject to quality data and data trust.”

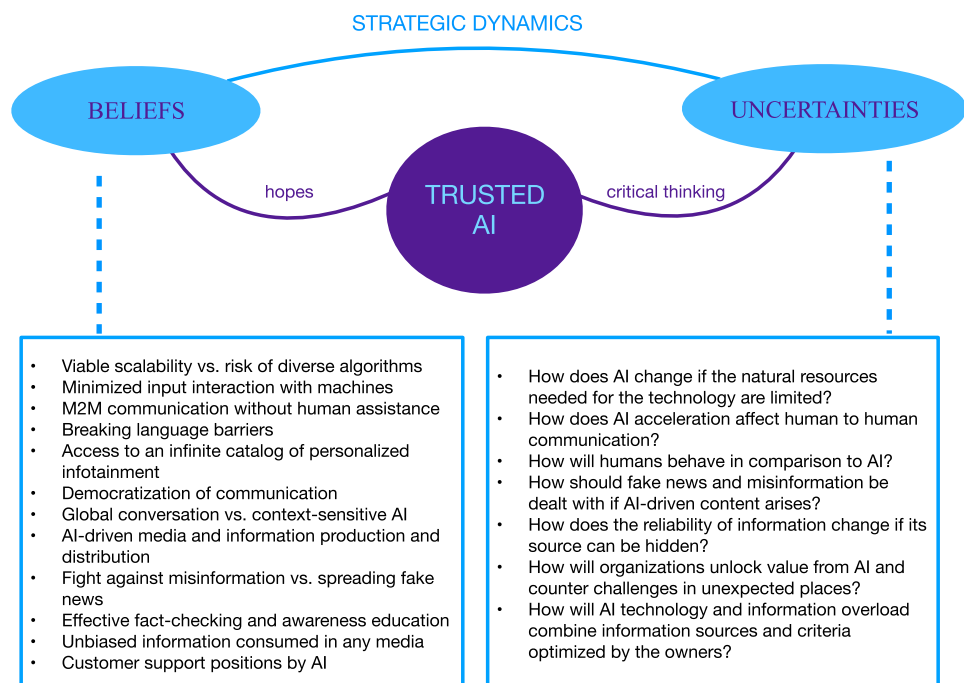
A female academic participant in knowledge management and tech consulting, Europe

The glasses model, therefore, underscores responsible usage and societal well-being for the distant future over profit maximization or political control. Although optimism is essential for the model, it also acknowledges potential negative impacts, emphasizing a balanced and realistic approach. Therefore, strategic dynamics in the trusted AI model represent a value-laden horizon for comprehensive policy-making, responsible business practice, and risk management.

The Glasses Model of AI Trust can be used to understand better the reasoning and rationale behind the development of further AI services within the ICM field, especially when reliability or ethical issues arise. Beliefs prioritize available opportunities by AI technology, and uncertainties clarify the direction leading to priorities along with appropriate questions. Examining the vertical results accordingly, an extended model summarizes the mostly techno-optimist beliefs and subsequent critical questions in cases of uncertainties (Fig. 2).

Most respondents emphasized how responsible thinking increases in the case of black box technology, primarily if AI impacts free will, privacy, power over people’s ability to make choices, think and interact effectively with others. Respondents argued along the lines of vision-focused questions and strategic proposals when broad questions were

Fig. 2 Glasses Model of AI Trust with a specification for ICM



raised, such as “what kind of a society do we want to be?” We also found a broad perspective with a concept of interconnectedness (Herbrechter et al. 2022) of how humans, machines and nature can connect in a sophisticated manner: if these ecosystems can learn from each other, it may lead to a less biased or perhaps even an unbiased future. Beliefs define the necessary inputs and scalability for the technology and the positive imaginations of democratized, less-fake, unbiased communication, information production and entertainment — resulting in more significant judgment, better interpretation, and creativity. Such positive beliefs and the identification of potential benefits were much more prevalent than any uncertainties among our study participants — which may have something to do with the self-selecting nature of our sampling method.

7 Discussion and conclusion

The paper aimed to present research in current and future AI-driven ICM process, surveying experts working in related fields and representing all world regions. Recalling the original assumptions, the experts have agreed on the first two:

- AI transforms ICM systems in several ways (Ellis and Tucker 2020; Guzman and Lewis 2020; Chan-Olmsted 2019; Gunkel 2019; Fletcher 2018; Waisbord 2018), and
- multifarious benefits are available from cost-effective operation to productive work (Mustak et al. 2023; Georgieva et al. 2022; Preu et al. 2022; Wirtz 2020).

The third assumption focused only on fake media, systemic bias and the possibility of abuses, given the context of profound uncertainties regarding AI (Borsci et al. 2023; Feher and Veres 2022; Rawat and Vadivu 2022; Holton and Boyd 2021; Bell 2018). However, this particular survey showed more comprehensive dynamics of visions strongly determined by participants’ personal beliefs and negotiation of uncertainties, which we modeled as a pair of glasses on a fulcrum — a *Glasses Model of AI Trust*.

The model is the key contribution of this paper to highlight emerging perspectives of AI among people who work in the field and who are responsible for implementing these technologies. The model allows for the analysis of abstract concepts, such as how an interconnected human–nature–technology ecosystem can improve AI, or how AI developments can fight against AI misuse. Thus, the model can be extended to different disciplines and industries. This model may contribute to testing specialized applications, future planning or evaluating other AI innovations. Policy agendas, developmental trajectories, and academic research are also invited to test the model and discuss the results. It seems clear that while these experts are keenly

aware of the pitfalls and problems of AI, their own beliefs and roles mitigate any profound concerns, as some participants went so far as to suggest that, in the future, AI will be the best defense against any problems caused or exacerbated by AI.

Last but not least, the fourth assumption was not accepted. According to the respondents, fundamental human/social values will not change. Although semi-truth news and demystifying the media process will be a new norm, such processes were not seen as directly affecting democratic values. Participants put much faith in new generations to continue to build balance and trust in emerging AI systems, which would preserve key social values.

To sum up, the experts linked their visions to the present and general AI discourses without abstract or specific associations. Two primary factors can result in this finding. First, their professional routines and responsibilities often lead them to view the distant future as a simple extension of the present. Second, they can be in professional near-sightedness of trending knowledge of AI, especially if they do not exclusively see themselves as responsible for the future directions. Both of these factors might be the reason behind the (over-)optimism as well.

The over-optimistic scenarios express hopes and beliefs, passing the responsibility on to future generations, technology and nature-inspiration. Although we did not set out to critique or question our study participants about their sense-making process in technology, techno-optimism holds risks and may fail to deal with worst-case scenarios (Del Rosso 2014).

Some themes do not have enough data to support them or the data are too diverse (Braun and Clarke 2006). Also, generative AI (van Dis et al. 2023) was not much hyped during the sampling process. These topics are missing and call for further research to understand the field from more diverse perspectives, disciplines and expertise.

Unexpectedly, the participants rarely mentioned certain fundamental topics of the field, such as platformization, the Metaverse, mixed realities, machine learning, sentient technologies, and more. Possible reasons for this could be that (1) these are not necessarily the future which they expect, (2) they do not use these applications regularly, or (3) they are unaware of the black box technology behind these services. We also cannot exclude the possibility that considering profoundly controversial aspects of AI is in itself problematic for those whose livelihoods are connected with the implementation and further development of AI. Their roles and commitment to their work may cause them to envision more favorable future directions, leading to a mostly optimist and partly pragmatist approach, according to Makridakis’s typology in AI scenarios (2017).

To summarize the key result, participants shared a positive and even hopeful vision of an AI-driven ICM process.

While they were aware of urgent questions regarding the growing technological determinism in how these fields (and their work) develop, they chose to remain hopeful based on personal beliefs and a limited questioning of remaining (and emerging) uncertainties regarding AI. They believe new generations are learning from the lessons of AI so far. Thus, they are passing the responsibility to the next generation and to new AI developments.

Consequently, the experts are not overly concerned about negative future scenarios. Although with this approach they also stated that they and their generation are not the only ones responsible for future AI in ICM, their optimistic interpretations and visions make them aware of their role as technology ambassadors. Thus, the experts in this study are fulfilling their role of being trust mediators (Bodo 2021).

Recalling one of the key messages of the experts, AI is both media and an entity/actor. Accordingly, AI does not only drive ICM but is seen as capable of fighting the disadvantages or misuse of advanced technology. According to the participants, the AI-to-AI era is coming, but AI-driven ICM will not radically transform social structures or values in this process. Moreover, the ultimate goals emphasized, such as focusing on well-being, security and sustainability; or avoiding colonization, political misuse and growing inequalities, can be feasible. Especially if universal access to trustable AI and “nature listening” confirm or improve social values, a hopeful future is envisioned. It is a rather optimistic, AI-driven future, where the paradox between a firm conviction of the sustainability of human values and an overreliance on AI-to-AI systems to remedy any problems and uncertainties in the process remains a blind spot for these experts, insofar as our survey is concerned.

In conclusion, the conflict between reliable and fake systems returns in the context of beliefs and uncertainties, but a value-saturated, typically reliable, AI-dominant vision emerges. Experts believe that what they build and trust in becomes available along the lines of hopes and uncertainty-based critical thinking. This positions them as mediators and ambassadors of AI connecting the current trends to the distant future, investing in a so-called “trusted AI”, even though it remains unclear from their answers what this trust would be based on – other than on more AI. Furthermore, they tend to disregard the possibility of AI assuming their expert roles as an agent in a relatively brief timeframe, producing data- and algorithm-driven visions and more diverse scenarios. It might be simply cognitive dissonance.

Finally, it is necessary to mention that the sampling process was completed just before the announcement of ChatGPT and the hype around generative AI. Consequently, this was last-minute research to receive long text-based responses from human experts without assuming machine-generated support. In the future, such research will be harder to imagine. The consequences of this need to be incorporated into

the analytical work — particularly if this work will also be supported by AI. The next question concerns what AI future will be forecast by applying joint human-AI analysis.

8 Limits

First and foremost, the sample was rather small and less diverse in terms of engineering knowledge, even if engineering and computer science are the dominant areas of expertise in the scientific literature regarding the question of the process. One of the reasons for this was that significantly fewer experts defined themselves on LinkedIn as being active in ICM, combined with engineering or computer science. Another reason could be that the survey questions seemed less relevant to them. Third, the participants provided relatively less new information than the available knowledge we could gather from the literature and trade publications. This is surprising, considering that the responses came from all regions of the world and several academic and professional fields. However, this result had two advantages. On the one hand, it summarizes the currently available knowledge and its limits. On the other hand, it made a more conspicuous interpretation of the professional role and attitudes. It comprehensively affects the ICM process and the direction of developments in AI broadly.

Funding Open access funding provided by National University of Public Service. The Janos Bolyai Research Scholarship of the Hungarian Academy of Sciences supported the paper, grant number: BO/00045/19/9 jointly with the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, Bolyai +. Grant number: UNKP-22-5-NKE-87. European Union’s Horizon Europe Research and Innovation Programme–NGI Enrichers, Next Generation Internet Transatlantic Fellowship Programme is also a fund of the research project and publication. Grant number: 101070125.

Data availability The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amaral G, Guizzardi R, Guizzardi G, Mylopoulos J (2020) Ontology-based modeling and analysis of trustworthiness requirements: preliminary results. International conference on conceptual modeling. Springer, Cham, pp 342–352. https://doi.org/10.1007/978-3-030-62522-1_25
- Araujo T, Helberger N, Kruijckemeier S, De Vreese CH (2020) In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI Soc* 35:611–623. <https://doi.org/10.1007/s00146-019-00931-w>
- Bell G (2018) Decolonizing Artificial Intelligence, Fay Gale Lecture, University of Adelaide, September. www.assa.edu.au/event/fay-gale-lecture-2/. Consulted 25 Jan 2019
- Blanchett N, McKelvey F, Brin C (2022) Algorithms, platforms, and policy: the changing face of canadian news distribution. In: Meese J, Bannerman S (eds) *The algorithmic distribution of news*. Palgrave global media policy and business. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-87086-7_3
- Bodo B (2021) Mediated trust: a theoretical framework to address the trustworthiness of technological trust mediators. *New Media Soc* 23(9):2668–2690. <https://doi.org/10.1177/1461444820939>
- Borsci S, Lehtola VV, Nex F, Yang MY, Augustijn EW, Bagheriye L, Zurita-Milla R (2023) Embedding artificial intelligence in society: looking beyond the EU AI master plan using the culture cycle. *AI Soc* 38:1465–1484. <https://doi.org/10.1007/s00146-021-01383-x>
- Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qual Res Psychol* 3:77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brevini B, Pasquale F (2020) Revisiting the Black Box Society by rethinking the political economy of big data. *Big Data Soc* 7(2):2053951720935146. <https://doi.org/10.1177/205395172093>
- Chan-Olmsted SM (2019) A review of artificial intelligence adoptions in the media industry. *JMM Int J Media Manag* 21(3–4):193–215. <https://doi.org/10.1080/14241277.2019.1695619>
- Couldry N, Mejias UA (2019) *The costs of connection*. Stanford University Press
- Del Rosso J (2014) Textuality and the social organization of denial: Abu Ghraib, Guantanamo, and the meanings of US interrogation policies. *Sociological forum*. Wiley Subscription Services, Cham, pp 52–74. <https://doi.org/10.1111/sof.12069>
- Ellis D, Tucker I (2020) *Emotion in the digital age: Technologies, data and psychosocial life*. Routledge. <https://doi.org/10.4324/9781315108322>
- Feher K (2020) Trends and business models of new-smart-AI (NSAI) media. In: 2020 13th CMI Conference on Cybersecurity and Privacy (CMI)-Digital Transformation-Potentials and Challenges (51275). IEEE, pp 1–6
- Feher K (2022) *Emotion artificial intelligence: Deep mediated and machine-reflected self-emotions. Mediatization of emotional life*. Routledge, pp 41–55
- Feher K, Katona AI (2021) Fifteen shadows of socio-cultural AI: a systematic review and future perspectives. *Futures* 132:102817. <https://doi.org/10.1016/j.futures.2021.102817>
- Feher K, Veres Z (2022) Trends, risks and potential cooperations in the AI development market: expectations of the Hungarian investors and developers in an international context. *Int J Sociol Soc Policy* 43(1/2):107–125. <https://doi.org/10.1108/IJSSP-08-2021-0205>
- Fletcher J (2018) Deepfakes, artificial intelligence, and some kind of dystopia: The new faces of online post-fact performance. *Theatre J* 70(4):455–471. <https://doi.org/10.1353/tj.2018.0097>
- Fossa F (2023) Data-driven privacy, surveillance, manipulation. *Ethics of driving automation: artificial agency and human values*. Springer Nature, Cham, pp 41–64
- Fuchs C (2009) Information and communication technologies and society: a contribution to the critique of the political economy of the internet. *Eur J Commun* 24(1):69–87
- Georgieva I, Timan T, Hoekstra, M (2022) Regulatory divergences in the draft AI act. Differences in public and private sector obligations. European Parliamentary Research Service. Scientific Foresight Unit (STOA) PE 729.507–May 2022. <https://doi.org/10.2861/69586>
- Glikson E, Asscher O (2023) AI-mediated apology in a multilingual work context: implications for perceived authenticity and willingness to forgive. *Comput Hum Behav* 140:107592. <https://doi.org/10.1016/j.chb.2022.107592>
- Gross PL, Buchanan N, Sané S (2019) Blue skies in the making: air quality action plans and urban imaginaries in London, Hong Kong, and San Francisco. *Energy Res Soc Sci* 48:85–95. <https://doi.org/10.1016/j.erss.2018.09.019>
- Gunkel DJ (2019) The medium of truth: media studies in the post-truth era. *Rev Commun* 19(4):309–323. <https://doi.org/10.1080/15358593.2019.1667015>
- Guzman AL, Lewis SC (2020) Artificial intelligence and communication: a human-machine communication research agenda. *New Media Soc* 22(1):70–86. <https://doi.org/10.1177/146144481985>
- Guzman AL, McEwen R, Jones S (2023) *The SAGE handbook of human-machine communication*. SAGE Publications Limited. <https://doi.org/10.4135/9781529782783>
- Hancock JT, Naaman M, Levy K (2020) AI-mediated communication: Definition, research agenda, and ethical considerations. *J Comput-Mediat Comm* 25(1):89–100. <https://doi.org/10.1093/jcmc/zmz022>
- Hartmann K, Giles K (2020) May) The next generation of cyber-enabled information warfare. *Int Conf Cyber Confl (CyCon)* 1300:233–250. <https://doi.org/10.23919/CyCon49761.2020.9131716>
- Hepp A (2020) *Deep mediatization*. Routledge, New York. <https://doi.org/10.4324/9781351064903>
- Herbrechter S, Callus I, Rossini M, Grech M, de Bruin-Molé M, Müller CJ (2022) *Critical posthumanism: an overview*. Palgrave. <https://doi.org/10.1007/978-3-030-42681-1>
- Hight C (2022) Deepfakes and documentary practice in an age of misinformation. *Continuum* 36(3):393–410. <https://doi.org/10.1080/10304312.2021.2003756>
- Holton R, Boyd R (2021) Where are the people? What are they doing? Why are they doing it? (Mindell) situating artificial intelligence within a socio-technical framework. *J Sociol* 57(2):179–195. <https://doi.org/10.1177/1440783319873046>
- Hui Y (2021) On the limit of artificial intelligence. *Philos Today* 65(2):339–357. <https://doi.org/10.5840/philtoday202149392>
- Illia L, Colleoni E, Zygliopoulos S (2023) Ethical implications of text generation in the age of artificial intelligence. *Bus Ethics Environ* 32(1):201–210. <https://doi.org/10.1111/beer.12479>
- Jackson D, Latham A (2022) Talk to The Ghost: the Storybox methodology for faster development of storytelling chatbots. *Expert Syst* 190:116223. <https://doi.org/10.1016/j.eswa.2021.116223>
- Jansen H (2010) The logic of qualitative survey research and its position in the field of social research methods. *Forum Qual* 11(2):11. <https://doi.org/10.17169/fqs-11.2.1450>
- Kemp S (2022) Digital 2022 global overview report. We are social and hootsuite. <https://wearesocial.com/uk/blog/2022/01/digital-2022-another-year-of-bumper-growth-2/>

- Kovtun V, Izonin I, Gregus M (2022) Model of functioning of the centralized wireless information ecosystem focused on multimedia streaming. *Egypt Inform J* 23(4):89–96. <https://doi.org/10.1016/j.eij.2022.06.009>
- Lin S, Hilton J, Evans O (2021) Truthfulqa: measuring how models mimic human falsehoods. ArXiv Preprint. <https://doi.org/10.48550/arXiv.2109.07958>
- Makridakis S (2017) The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures* 90:46–60. <https://doi.org/10.1016/j.futures.2017.03.006>
- Mansell R, Steinmueller WE (2020) *Advanced introduction to platform economics*. Edward Elgar Publishing, Cheltenham
- Mustak M, Salminen J, Mäntymäki M, Rahman A, Dwivedi YK (2023) Deepfakes: deceptions, mitigations, and opportunities. *J Bus Res* 154:113368. <https://doi.org/10.1016/j.jbusres.2022.113368>
- Newlands G (2021) Lifting the curtain: strategic visibility of human labour in AI-as-a-service. *Big Data Soc* 8(1):20539517211016024. <https://doi.org/10.1177/20539517211016026>
- Pavlik JV (2023) Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education. *J Mass Commun Educ* 78(1):84–93. <https://doi.org/10.1177/10776958221149577>
- Pierson J, Kerr A, Robinson C, Fanni R, Steinkogler V, Milan S, Zampedri G (2023) Governing artificial intelligence in the media and communications sector. *Internet Policy Rev*. <https://doi.org/10.14763/2023.1.1683>
- Preu E, Jackson M, Choudhury N (2022) Perception vs reality: understanding and evaluating the impact of synthetic image deepfakes over college students. *IEEE Thirteen Ann Ubiquitous Comput Electr Mob Commun Conf (UEMCON)*. <https://doi.org/10.1109/UEMCON54665.2022.9965697>
- Rassameeroj I, Wu SF (2019) Reverse engineering of content delivery algorithms for social media systems. *Sixth Int Conf Soc Netw Anal Manag Secur (SNAMS)*. <https://doi.org/10.1109/SNAMS.2019.8931859>
- Rawat S, Vadivu G (2022) Media bias detection using sentimental analysis and clustering algorithms. *Proceedings of international conference on deep learning computing and intelligence*. Springer, Singapore, pp 485–494. https://doi.org/10.1007/978-981-16-5652-1_43
- Ryan M (2020) In AI we trust: ethics, artificial intelligence, and reliability. *Sci Eng Ethics* 26(5):2749–2767. <https://doi.org/10.1007/s11948-020-00228-y>
- Stewart MC, Arnold CL (2018) Defining social listening: recognizing an emerging dimension of listening. *Int J List* 32(2):85–100
- Swiatek L, Galloway C, Vujnovic M, Kruckeberg D (2022) Artificial intelligence and changing ethical landscapes in social media and computer-mediated communication: considering the role of communication professionals. *The Emerald handbook of computer-mediated communication and social media*. Emerald Publishing Limited, Cham, pp 653–670. <https://doi.org/10.1108/978-1-80071-597-420221038>
- Targowski A (2019) The element-based method of civilization study. *Comp Civiliz Rev* 81(81):6. <https://doi.org/10.1080/10904018.2017.1330656>
- Toy T (2023) Transparency in AI. *AI Soc*. <https://doi.org/10.1007/s00146-023-01786-y>
- Trattner C, Jannach D, Motta E, Costera Meijer I, Diakopoulos N, Elahi M, Moe H (2022) Responsible media technology and AI: challenges and research directions. *AI Ethics* 2(4):585–594. <https://doi.org/10.1007/s43681-021-00126-4>
- Van Dijck J (2021) Seeing the forest for the trees: Visualizing platformization and its governance. *New Media Soc* 23(9):2801–2819. <https://doi.org/10.1177/1461444820940293>
- van Dis EA, Bollen J, Zuidema W, van Rooij R, Bockting CL (2023) ChatGPT: five priorities for research. *Nature* 614(7947):224–226. <https://doi.org/10.1038/d41586-023-00288-7>
- Verbeek PP (2015) Cover story beyond interaction: a short introduction to mediation theory. *Interactions* 22(3):26–31. <https://doi.org/10.1145/2751314>
- von Eschenbach WJ (2021) Transparency and the black box problem: why we do not trust AI. *Philos Technol* 34(4):1607–1622. <https://doi.org/10.1007/s13347-021-00477-0>
- Waisbord S (2018) Truth is what happens to news: On journalism, fake news, and post-truth. *J Stud* 19(13):1866–1878. <https://doi.org/10.1080/1461670X.2018.1492881>
- Wirtz J (2020) Organizational ambidexterity: cost-effective service excellence, service robots, and artificial intelligence. *Organ Dyn* 49(3):1–9. <https://doi.org/10.1016/j.orgdyn.2019.04.005>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.