



Public accountability and regulatory expectations for AI in journalism: qualitative evidence from focus groups with Dutch citizens

Sophie Morosoli¹ · Laurens Naudts¹ · Hannes Cools¹ · Karthikeya Venkatraj² · Natali Helberger¹ · Claes de Vreese¹

Received: 1 May 2025 / Accepted: 25 August 2025
© The Author(s) 2025

Abstract

As artificial intelligence (AI) continues to reshape society, its integration into journalism raises critical questions about transparency, accountability, and public trust. Existing AI regulations have largely been developed without meaningful public input, prompting concerns about whether current governance approaches adequately address societal expectations. This study investigates the expectations and concerns of Dutch citizens regarding mandatory AI disclosures in journalism through three focus groups ($N=21$). We aimed for a broad sample of participants to ensure diversity in terms of age, gender, and education level. Key questions measured the main concerns about AI-generated content, why participants want to know if they are interacting with AI-generated content and which rights individuals would like to have in this context. The results reveal a preference for participatory regulatory processes and standardized transparency measures, such as the disclosure of sources. The results further underscore the wish to be able to hold news organizations and individual AI users accountable when regulations are breached. The findings can inform news professionals and regulators alike, for example, in the context of the implementation of the AI transparency obligations in the European AI Act.

Keywords Artificial intelligence · Journalism · Governance · Accountability · Individual rights

1 Introduction

Over the past years, we have seen a vivid debate about how artificial intelligence (AI) is transforming society, and the many potential risks like threats to privacy, fears of job displacement, the amplification of disinformation, as well as what kind of ethical frameworks and governance solutions

are needed to protect individuals and society. Only: where are the citizens in this debate? Although AI has become a significant part of everyday life, regulatory frameworks are often developed without meaningful input from the public (König 2023). This exclusion poses a risk of alienating citizens and creating a gap between regulatory decisions and societal needs (Strycharz et al. 2020). This disconnect is particularly noticeable in fields where trust and transparency are crucial (Weymouth 2015). One example of this is journalism, the focus of this study. To render transparency regulations in journalism effective, an understanding of what kind of information people need to trust news, including the use of AI *in* news, is required.

For at least five years, AI has become an integral part of the news production process (Simon 2024). From automated news writing and data-driven investigative reporting to algorithmic content personalization, AI technologies are widely employed to enhance efficiency and reach (Diakopoulos et al. 2024). While these innovations offer numerous benefits, they also raise important questions about transparency and accountability. Citizens increasingly express a desire to understand when and where AI has been used in the production of news (Piasecki et al. 2024; Toff and Simon

✉ Sophie Morosoli
s.v.morosoli@uva.nl

Laurens Naudts
l.p.a.naudts@uva.nl

Hannes Cools
h.cools@uva.nl

Karthikeya Venkatraj
karthikeya.venkatraj@cw.nl

Natali Helberger
n.helberger@uva.nl

Claes de Vreese
c.h.devreese@uva.nl

¹ University of Amsterdam, Amsterdam, Netherlands

² Centrum Wiskunde & Informatica, Amsterdam, Netherlands

2023). This wish for transparency, in turn, can have direct consequences for the trust relationship individuals have with news organizations.

Despite these pressing concerns, existing regulatory approaches often fail to address the specific needs of citizens regarding AI disclosures. For instance, the European Union's AI Act, one of the most ambitious legislative efforts to regulate AI so far, does not impose transparency obligations on news organizations about their use of AI as long as specific content has undergone sufficient editorial oversight.¹ And this despite empirical insights highlighting that individuals find being informed that AI has been used in the news production important (Piasecki et al., 2024). Furthermore, the AI Act's transparency obligations have not been informed by empirical research into concrete information needs of users (Piasecki et al. 2024), and the same is true for many ethical journalistic AI guidelines that emerge and put transparency central (Cools and Diakopoulos 2024). Examples like these highlight a significant disconnect between existing regulations and the expectations of citizens, who view transparency as essential for having their information needs met and holding news organizations accountable.

To address the current lack of the involvement of the public in the development of AI governance and the discussion of what rights users should have in relation to the use of AI, this study examines individuals' perspectives on the regulation of AI disclosures in journalism. By conducting focus groups with citizens of the Netherlands, this study offers a comprehensive, in-depth understanding of individuals' concerns, expectations, and preferences regarding transparency and accountability related to the use of AI in journalism. This study, therefore, aims to answer the following research questions:

What are individuals' views on regulating AI disclosures in journalism, and how do they envision these regulations?

If potential regulations are violated, who do individuals think should be held accountable, and how?

2 Individuals' concerns about AI & individuals' need for rights

To better understand what regulatory needs citizens have, it is important to take a step back and look at the concerns they might have. These concerns, in turn, might prompt specific needs for laws or ethical guidelines. We observe that AI triggers a multitude of far-reaching societal concerns. For instance, findings from *The Ada Lovelace* and *The Alan*

Turing Institute show that British citizens are mostly concerned about AI when it comes to autonomous weapons and advanced robotics (Modhvardia 2023). Even though British adults see advantages in using AI, for example, for the detection of cancer, they also express worry about relying too heavily on AI technologies, the lack of transparency and accountability connected to the deployment of AI, and the collecting of personal data against one's will (Modhvardia 2023). In the U.S., a majority of citizens feel more concerned than excited about the influence AI has on daily life (Tyson and Kikuchi 2023). Similar to British people, Americans are especially worried about the fact that AI will have a negative impact on their privacy, followed by companies' inability to provide good customer service (Tyson and Kikuchi 2023). In the Netherlands, where this study is situated, trust in AI is low compared to other countries, and the biggest concern Dutch people have concerning AI is AI-generated disinformation and general abuse [Department for the Coordination of Algorithmic oversight (DCA), 2024]. Furthermore, citizens have little-to-no trust that automatic decision-making systems can accurately determine social welfare, the news one sees, and cancer diagnoses (AlgoSoc 2024).

These societal concerns also relate to the focus of this study—the application of AI in journalism. Zooming in on these concerns, previous research suggests that citizens worry about multiple things. The Digital News Report of the Reuters Institute highlights that individuals refuse full automation in the news production process in all countries included in the report. Individuals see the most danger in the use of AI in journalism when it is used to create images or videos that seem to be real (Newman et al. 2024). Evidence from China shows that Chinese citizens are mostly concerned about echo chambers and privacy violations connected to the use of AI in journalism (Sun et al. 2022).

Directly connected to these concerns, previous research has identified that individuals feel the need for rights or legal protection when it comes to AI. Over 60% of British citizens wish for laws regulating AI technologies, and when being asked who should be responsible for the safe use of AI, 41% of the participants indicated that they would want an independent regulator to do so (Modhvardia 2023). Related to this, the participants further stressed that they want to better understand how AI makes decisions, and hence, explainability and transparency are of the essence among British citizens (Modhvardia 2023). In Germany, citizens support overall strong or moderate regulation of AI, considering the influence of AI on the environment and personal autonomy. Hence, the wish for regulation is present if individuals feel AI could induce long-term societal harm. Similarly to British citizens, individuals also put their focus on regulating or providing transparency, especially connected to AI's impact on personal autonomy (König 2023).

¹ Article 40 (4) AI Act.

However, despite the far-reaching consequences of AI on society and strong individual concerns, AI regulations are often created by policymakers without the meaningful involvement of citizens (OECD 2024). Even though the success and approval of such policies are dependent on individuals' preferences (König 2023; Wong et al. 2022). AI regulation without citizen input risks overlooking the needs and concerns of the people most affected by these technologies, but also risks designing regulations that do not work for people. Wong et al. (2022) suggest participatory governance for tackling AI governance, which means that citizens are involved in the policymaking process through deliberative practices.

3 Regulating AI disclosures²

From a regulatory perspective, the EU has positioned itself as a frontrunner in human-centered technology legislation, in which citizens are ensured “an adequate level of transparency about the use of algorithms and artificial intelligence,” empowered in their use of AI, and informed concerning their interaction with AI (European Parliament, the Council and Commission, 2022). The recently adopted AI Act reflects those lofty ambitions, and in so doing, explicitly addresses the concerns citizens might have regarding the uptake of AI in news and media environments. Most notably, recital 133 AI Act³ cautions against generative AI-systems in particular: their capacity to generate large amounts of seemingly human and authentic content threatens the “integrity and trust in the information ecosystem, raising new risks of misinformation and manipulation at scale.” Unfortunately, the AI Act risks becoming a paper tiger in the context of automated news production due to several interrelated limitations. First, the AI Act relies on transparency and disclosure obligations as core protective mechanisms, but fails to leverage the instrumental function this information can hold for empowering citizens. Moreover, the media sector benefits from disclosure exemptions. Second, beyond transparency, citizens have no concrete rights to directly contest the use of AI in news environments or request alternative, non-AI-generated news.

3.1 Transparency in the AI act

Given the “wide availability and increasing capability of [generative AI systems]” and the epistemic risks associated with people's inability to “distinguish human-generated and

authentic content” from artificial and inauthentic content (Recital 133 AI Act), the EU lawmaker requires that citizens are informed of their interaction with AI-systems and AI-generated content (Article 50 AI Act). Transparency duties also affect the design and operations of AI-systems. For example, providers of systems that can generate synthetic audio, image, video, or text content should ensure that the output of these systems can be “marked [and detectable] as artificially generated or manipulated (Article 50, para 2 AI Act)”. Though certainly necessary, in their current form, the AI Act's transparency and disclosure duties remain an insufficient precondition toward citizen empowerment (Beckett 2023; Piasecki et al. 2024). Consider the threats of misinformation and manipulation that Article 50 AI Act sought to address: First, informing citizens about their interactions or consumption of artificial content, even in the case of deep-fakes, is one thing, but imparting knowledge to them regarding the risks such content might entail is another (see also: De Andrade et al. 2023). Simply informing them that content has been AI generated does not signal to users whether that content is trustworthy or not, and without sufficient digital literacy skills, citizens might not even know that artificial content can be prejudiced. Second, artificial content need not be misleading, nor manipulative. Media organizations currently experiment with AI to assess how these novel technologies can responsibly support the news production and communication life cycle (Beckett 2023). Their goal is to preserve values of journalistic integrity within and through AI. Finally, given the informational risks AI entails and the trust placed in transparency, it is rather puzzling that the journalism industry, whose watchdog function is critical for a well-functioning democracy, has been largely exempt from these obligations. In principle, citizens must be notified when they consume textual, public interest news content generated or manipulated by AI. However, this “obligation shall not apply [...] where the AI-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content”. This carve-out reflects a general reluctance against regulation to protect the media from government intervention and censorship, relying instead on self-regulation (see also below). Having said that, without such self-regulation, there would be a real risk that individuals would not be told they consume AI-generated news content. At the same time, in the case that people would be informed, telling people a particular piece of information is artificial does little to empower citizens in distinguishing truthful from manipulative and deceptive content.

² This section draws upon prior work of the authors, some of which was translated from publications which first appeared in their native language, and in particular: anonymized for peer review.

³ <https://artificialintelligenceact.eu/recital/133/>

3.2 Beyond transparency

Transparency can, of course, serve as a sound foundation for citizen empowerment. To do so, however, two questions are critical. First, does transparency help citizens in their follow-up actions vis-à-vis the use of AI in news media settings? This is a question of what information enables citizens to do. As aforementioned, merely informing people about their interaction with AI gives little insight into the actual risks people are exposed to. To give them better insights into potential risks, they might need additional or different information, for instance, insights into the data or normative assumptions upon which the model was trained, whether the content has been checked by a human, or what the likelihood of inaccuracies is. While this would not necessarily enable *all* citizens to make a reliable assessment regarding the flaws AI-generated content might exhibit, it could enable some (tech-savvy) citizens and interest groups to scrutinize and contest these technologies. Likewise, transparency can be used to inform citizens not only of the risks they are subject to but also of the rights they might have to counter such exposure, and in so doing, empower citizens in the exercise of these rights (e.g., Kaminski 2019; Naudts et al., n.d.; Varošaneč 2022).

Second, can citizens effectively pursue follow-up actions? If the goal of transparency obligations is to empower people to make informed choices or defend their rights, citizens also need the rights to do so in the first place. In this context, citizen rights are particularly important as they give people a legal mandate that legitimizes their claims against the use of AI. Notwithstanding other legal instruments, the AI Act itself offers a few actionable rights to citizens. Article 85 in the AI Act stipulates that every natural or legal person can submit a complaint to the relevant market surveillance authority in case they believe the regulation has been infringed upon. This would require that people know that a piece of content was artificial but not labeled and decide to bring this failure to comply before the relevant authority—an unlikely scenario. Citizens also have no right to demand additional information or a non-AI-generated alternative, and no right to complain directly against those who use AI, such as the news organization's managerial or editorial board. Article 86 of the AI Act does offer citizens a right to an explanation in the case a decision is made by the deployer of an AI system based on the output of that AI system. Yet, this is unlikely to apply within the context of AI-generated news production. First, it is highly uncertain under what conditions news content could constitute a decision in the aforementioned legal sense. Second, the right to an explanation only applies to high-risk systems. However,

the media sectors are typically not considered high-risk contexts under the AI Act.⁴

In other words, the AI Act does too little to hold news content providers accountable and empower news consumers to make truly informed decisions. In a previous study, however, the authors found that Dutch citizens would like to receive additional tools when navigating their news environments (Piasecki et al. 2024). Three follow-up actions in particular stood out: the ability to a) filter AI from non-AI news, b) contest AI-generated information, and c) report biased news provision. At the same time, Piasecki and colleagues argue that “simply informing people about the fact that a text has been automatically generated or manipulated by AI gives them little cues about how to interpret and assess the article.” (Piasecki et al. 2024, p. 15). Rather than enable citizens to co-construct their digital news environment, citizens are instead left in the dark with a few alternatives to choose from (Naudts et al. 2025).

3.3 Transparency in newsrooms

For the past decades, the norm of transparency has been institutionalized to make journalism more accountable, credible, trustworthy, and legitimate in an increasingly digital, interactive, and automated news environment (Diakopoulos and Koliska 2017; Karlsson 2010, 2021). Due to a steady decline in media credibility and trust (Newman et al. 2024) and greater adoption of digital technologies, there has been a growing emphasis on this ethical norm, as well as how transparency is practiced within professional journalism (Cools and Koliska 2024).

With the emergence of generative AI tools in 2022, the discussion of transparency as a norm and a practice became front and center in newsrooms around the world, as these tools were, and are able to gather, produce, verify, and distribute content (Diakopoulos et al. 2024). Similarly, these discussions on transparency, as well as reflections on ethical AI use, were relatively soon crystallized into charters, principles, or guideline documents that were proposed by news organizations (Cools and Diakopoulos 2024; Dodds et al. 2024).

Many of these discussions were happening in a longer tradition of self-regulation on which the journalistic profession has relied on (Becker et al. 2025). Generally, self-regulation rests on two pillars with (A) many news organizations regulate themselves by setting up non-governmental press councils spanning the industry, and by (B) proposing guidelines on their own organizational level (Cools &

⁴ The regulation primarily targets the use of AI-systems in high-risk contexts, such as hiring or social welfare algorithms; see Annex III AI Act.

Diakopoulos 2023). A similar development was witnessed during the emergence of, and the institutionalization of the use of social media platforms by news organizations who started to draft their own guidelines, as well as proposing more higher level principles in overarching press industry charters (Opgenhaffen and d’Haenens 2015).

Just like the social media guidelines and charters, not a lot of homogeneity across organizations, as well as ambivalence, has also been found in guidelines around the use of AI (Becker et al. 2025). Overall, the norms like ‘human supervision’, ‘transparency’, and ‘bias’ that are outlined in these documents are similar, but this has not necessarily led to an operationalization of these guidelines, which has been described in literature as the ‘Principles to Practices Gap’. Generally, this concept refers to the gap between organizational intent and internal understanding of operationalizing this effectively across teams and workflows on the ground. In 2025, scholars have criticized that the guidelines of these news organizations remain vague, which indeed leads to a disconnect between these principles and practices (Chen et al. 2025; Sadek et al. 2025). For example, the Dutch Press Agency (Algemeen Nederlands Persbureau) describes human supervision as follows: “In our production chain, we stick to the already established line human > machine > human where thinking and decision-making begins and ends with humans”. Similarly, for transparency, The Guardian says that when they use generative AI, they will be open “with their readers when they do this”.

This vagueness of guidelines complicates the way in which news organizations communicate about their uses of AI with their audience. Over time, the news industry has been rarely regulated, and this tradition has also led to a ‘soft law’ approach where a lot of the guidelines, charters, and principles they come up with are left to interpretation, or are simply not addressed (Paik 2023).

4 Method

To gain an in-depth understanding of individuals’ needs for rights and regulatory entitlement in the context of AI disclosures, we conducted three focus groups with Dutch citizens. The participants were recruited by a local panel company (Bureau Fris), and the focus groups were conducted in person at the authors’ university. While focus groups are not designed to produce statistically representative samples, we aimed to approximate a diverse cross-section of participants to ensure a broad range of perspectives relevant to the topic of this study. Each group consisted of seven individuals who were diverse in age, gender, and education levels. The youngest participant was 24 years old, and the oldest was 63 years old. Eleven participants identified as female and ten as male. Eight individuals had a lower education level,

eight a medium and five a high level of education. Four individuals were students, three were without an occupation, and 14 were part of the workforce. The focus groups lasted approximately one hour and were conducted by the authors in the summer of 2024. In each focus group, two authors were present—one communication scientist and one legal scholar—with one of them leading the interview, while the other took notes. Focus groups have multiple advantages. One such benefit is the observed interaction between participants, which closely resembles everyday conversations and, as a result, enhances external validity (Peek and Fothergill 2009). Moreover, the absence of structured interview or survey constraints enables participants to articulate their opinions freely, thus facilitating the exploration of a broader range of topics and enhancing the discernment of their underlying reasoning and thought processes (Morgan 1996).

This study is situated in the Netherlands, which is a particularly interesting case for studying artificial intelligence and its governance. Already in 2019, the Dutch Government released a strategic action plan for AI, in which the government describes strategies for how the Netherlands will tackle AI (European Commission 2021; Ministerie van Binnenlandse Zaken en Koninkrijksrelaties 2024). For instance, the government encourages the development and innovation of AI to address societal challenges such as climate change and health care systems under stress. Furthermore, the government announced that it would be investing in research in AI, data access, and AI-related skills development (European Commission 2021; Ministerie van Binnenlandse Zaken en Koninkrijksrelaties 2024). Hence, the Netherlands is a country where the development and governance of AI are political focal points. Legally, as an EU member state, the Netherlands must comply with the recently introduced AI Act. At the micro-level, around 30% to 40% of Dutch citizens indicate that they are familiar with AI-related terms, such as ChatGPT and algorithms. However, 70% of the Dutch people do not (yet) use generative AI tools (AlgoSoc 2024).

The set-up of the focus group was as follows: All focus groups started with the interviewer asking the participants to write down the three main concerns they have about AI-generated content. These concerns were then ranked by the whole group from least to most concerning. After this first part of the focus group, the participants were presented with multiple news headers (including images) and news articles that were generated by AI. The participants were informed about this fact and were given as much time as they needed to look at all of them. The headers included different news topics, such as politics, sports, and crime. They either only contained AI-generated text or a headline and an AI-generated image. Figure 1 shows two examples of the stimuli.

For this second part of the interviews, participants engaged in a plenary discussion about their information needs, expectation management, need for rights, and the

Klimaatactivist Gearresteerd voor Bevuilen van Monet-schilderij in Parijs



Een klimaatactivist is zaterdag gearresteerd nadat zij Claude Monet's schilderij "Coquelicots" uit 1873 in het Musée d'Orsay in Parijs had beklad. De activiste, lid van de milieuorganisatie Riposte Alimentaire, wilde de aandacht vestigen op de klimaatcrisis door een bloedrode poster over het kunstwerk te plakken.

Het incident vond plaats op zaterdag toen een vrouwelijke activist van Riposte Alimentaire een plakkerige poster met daarop een kaal, apocalyptisch landschap over Monet's "Coquelicots" bevestigde.

Zelensky's Luxe Leven en Misbruik van Amerikaanse Hulp Ontketenen Wereldwijde Controverse terwijl Jachtschandaal Oekraïne Opschrikt



Terwijl Oekraïne strijdt tegen Rusland, brengen onthullingen over luxe jachten en grootschalige corruptie het vertrouwen in de nationale leiding aan het wankelen.



Fig. 1 Examples of AI-generated news content used in focus groups

authenticity of AI-generated content. In this study, we focus on answers surrounding individuals' information and their regulatory needs and their responses regarding accountability regarding the use of AI in journalism.

The focus groups were recorded, transcribed, and translated by the four authors. To ensure that the identity of the participants is protected, we anonymized the data presented in this study using "Subject 1, Group 1" as a reference to a participant. For readability purposes and to enhance the clarity of the quoted materials, we edited filler words and stammering out of the presented quotes.

For the analysis, we adopted an inductive approach using thematic analysis to identify recurring patterns (Braun & Clarke 2013). A method that is known for its flexibility in terms of sample size, data collection, and theoretical flexibility (Clarke & Braun 2017). To do so, the authors first read the transcripts multiple times to become familiar with the content, and each author created an individual codebook using open coding, identifying themes and possible codes. Based on these codebooks, a consensus codebook was created with all authors present (see Appendix A). The authors then proceeded to analyze all the transcripts line-by-line together, discussing and coding each statement using the software Nvivo. Throughout the coding process, the codes and themes were thoroughly and regularly discussed by all authors to ensure validity. This "negotiated agreement" process was relied on to ensure intercoder reliability (Campbell et al. 2013).

5 Results

The results of this study are structured along three main themes and related codes identified through thematic analysis: regulation, information needs, and concerns about AI in journalism. Table 1 shows an overview of the themes and codes identified in this study. First, the theme 'regulation' identified a strong call for regulatory entitlement in the use of AI in journalism, including the protection against AI-related risks such as misinformation. Directly resulting from

Table 1 Overview of identified themes and codes

Regulation
Protection against risks
Regulatory entitlement
Accountability & responsibility
Information needs
Verification, regulation and oversight
Human oversight
Institutional reliability
Standards and guidelines
Concerns of AI in journalism
Informational concerns
Lack of accuracy and truthfulness (fake news)
Malintent/abuse
Polarization
Difficulties recognizing AI vs. non-AI information
No or less diversity and aims for inclusivity in AI
Lack of authenticity in AI
Distrust AI
Existential fear of AI
No human oversight/losing control

this is the third sub-theme ‘accountability and responsibility’, which describes individuals’ opinions about who should be held accountable or take responsibility in the context of AI in journalism. Connected to this is the second main theme, ‘**information needs**’, where participants emphasized the importance of verification, regulation, and oversight in the use of AI within journalism. A key concern was the need for human oversight to ensure that AI-generated content is reviewed and fact-checked before publication.

Participants also highlighted the value of institutional reliability, suggesting the introduction of labels to signal trustworthy AI use. Furthermore, we saw a need for guidelines on how AI is integrated into journalistic practices, aimed at safeguarding journalistic values. Finally, this study relies on the theme of ‘**concerns of AI in journalism**’. These are primarily centered around the negative impact of AI-generated content. For instance, fears about misinformation, manipulation, polarization, and the difficulty in distinguishing AI-generated from human-created content. These issues also brought up feelings of distrust and, in some cases, existential fears about AI, such as job displacement.

5.1 Regulation & information needs

First and foremost, we identify a clear need for regulation. This need is closely connected to the desire to be protected against existential risks related to artificial intelligence. Participants expressed a clear need for rights, or regulatory entitlement this way: *Subject 4 (Group 2): “I think we can say anyway and, for example, Elon Musk also said this once that if we continue at this rate now with the developments in AI then we are going way too far. And that, we shouldn’t want that as humanity. So I don’t think it’s too late now as a society to make laws and regulations about that now and with all of us to think about how do we envision this and how would we like to use that? And I think it’s very good that we’re going to make rules about that, and it’s not too late, I think”.* *Subject 2 (Group 3) adds: “I agree that you have to be sharp that we have to start now with indeed rules, and things like that so that it can’t take us over later on”.* It becomes apparent that the participants share some existential fears that go beyond the use of AI in journalism, and they wish for regulation that protects them against these dangers.

Looking at the impact of AI on journalism and the related need for regulations, we observe the following: *Subject 5 (Group 1): “Look, of course, you often have multiple sides, illuminating a truth from multiple perspectives, but AI who can also just bend the truth a.k.a make it up a bit more. So yes, I would like it if this is really regulated”.* *Subject 1 (Group 2) added: “But when it comes to news, I would say that good regulation should come. That the source is always behind as well. Whether you use it for shortening,*

fine, but those sources should be disclosed, it seems to me. That would be my expectation that good regulations will be made. But if you just let it go that way, you never know where it’s going to end up”. Another individual directly referred to a set of rules or regulation within the journalistic realm: *Subject 3 (Group 1): “And you indicated there is a kind of code of conduct for journalism, so then that should also be part of that, that it’s just the same rules for everyone and, indeed, if you use AI in journalism, that you’re just bound by certain rules, that you then know it may have been generated with AI but I know that certain rules are behind it”.* Thus, the participants express a need for more information when it comes to the use of AI in journalism, such as consistent source disclosure and transparent internal guidelines.

Some individuals went one step further and expressed that they even feel entitled to the truth or the protection of certain rights. As *Subject 2 (Group 1)* put it: *“You obviously want to read the truth. Only how nuances are just a little differently written down, and then a few words are different, then suddenly the story can be read very differently. So I can say: yes, I am entitled to good news, or the truth, but that can also be tweaked, of course”.* The participant not only expresses a right to truthful content but also acknowledges the interpretive nature of language and framing, especially in the context of AI-generated content.

However, this feeling of regulatory entitlement also triggered the participants to question who should regulate the use of AI in journalism. For instance: *Subject 5 (Group 2): “I don’t see-who should do that then. I mean: who is going to regulate that? I don’t get it”.* This uncertainty was reinforced by *Subject 2 (Group 2)*, who reflected on a hypothetical scenario about watermarking AI-generated media and then immediately questioned the feasibility of enforcement: *“If someone posts something then, that person is not going to... Suppose you say every AI movie for example must have a watermark, so you know it’s AI generated. Who is going to regulate that?”.*

In response to this regulatory uncertainty, several participants proposed institutional reliability as a solution, pointing to the need for recognizable, trustworthy symbols or certifications. For instance, *Subject 6,(Group 1): “Of course, it also has to be a, a recognized logo, you just see that in the normal world as well that there are a lot of fake logos in circulation and people are being misled anyway. And it has to be something official”.* Or *Subject 1 (Group 1): “Yes, so instead of it being ‘generated by AI’ that it says: this has an ‘AI seal of approval’ and it complies with these rules”.* These remarks reflect an ambivalence: while institutional disclosures are seen as necessary, participants were simultaneously aware of the limitations and vulnerabilities of such ideas, like fake labels.

In sum, participants expressed a complex yet consistent concern about the unregulated application of AI in

journalism. While fears about AI's broader societal impacts were evident, these concerns were translated into a wish for protective regulation, transparency, and reliable information. Participants expressed a regulatory entitlement, an ideal that society not only needs but deserves legal safeguards. Individuals proposed institutional reliability to realize and ensure, for example, the creation of recognizable disclosures such as seals of approval or official certifications to help recognize trustworthy (AI-generated) content. However, this idea was also met with skepticism, acknowledging that these types of disclosures are also vulnerable to manipulation. Moreover, these discussions hint toward participants identifying a false dichotomy between artificial and truthful content, and the dangers of enforcing what constitutes truthful information through laws.

5.2 Accountability & responsibility

Closely linked to the broad theme of regulation is the sub-theme of accountability and responsibility. For instance, when the (hypothetical) rights discussed above would be violated, multiple individuals indicated that they feel the need to have effective accountability mechanisms in place.

A key point raised by several participants was the importance of human accountability. Subject 3 (Group 3) stressed this by stating: *"If something is written by AI, you need to know that,"* underscoring a demand for transparency that enables attribution. Subject 2 (Group 3) extended this idea, arguing that the person using AI should bear ultimate responsibility, as they remain the final decision-maker. These reflections convey a consensus: while AI may assist in content creation, the ultimate responsibility lies with the person or institution who uses the technology.

Beyond accountability, some participants also raised the question of what should happen in case certain regulations have been ignored or circumvented. For example, Subject 1 (Group 1) asked: *"Should there be a fine on it? A sentence,"* while Subject 6 (Group 3) added: *"And as a matter of fact, punishable had you bypassed that anyway".* These remarks highlight a desire not only for rules but for mechanisms to ensure compliance, including potential punitive measures. This led to further reflection on what accountability could look like in a future where AI is part of every aspect of society. Subject 6 (Group 3) speculated: *"I can imagine that at some point you get a kind of court system for computers [...] that you can have the computer system on its own shut down".* Yet, this idea was met with disagreement by Subject 1 (Group 3), who pointed out: *"But that computer still belongs to someone, so I do think that whoever owns that computer is just responsible".* There were more futuristic suggestions, such as the notion that AI itself should be trained in rights and duties, as Subject 6 (Group 3) further

proposed: *"You should also start teaching AI about rights and duties".*

Focusing specifically on AI in journalism, participants emphasized the necessity of **human oversight and verification** as a critical safeguard. Subject 7 (Group 3) argued: *"That AI is controlled by a person. Not 100 percent of generative AI but a person should control for news."* This sentiment was echoed by Subject 5 (Group 1): *"You always have to do some kind of check and not just blindly throwing in a source. And think here: publish, enter, done."* Similarly, Subject 2 (Group 3) described a scenario where AI is used to support, but not replace, journalistic integrity: *"If you know, okay, he looked it up himself, scanned it in himself for example [...] he just accelerated it, so to speak."*

Looking at the specific case of the use of AI in journalism, which we focus on in this study, the main approach to accountability that the participants named was human oversight and verification: Subject 7 (Group 3) argued: *"That AI is controlled by a person. Not 100 percent of generative AI but a person should control for news".* This sentiment was echoed by Subject 5 (Group 1): *"In itself for a user, yes, and probably for the journalist too, but you always have to do some kind of check and not just blindly throwing in a source. And think here: publish, enter, done".* Similarly, Subject 2 (Group 3) described a scenario where AI is used to support, but not replace, journalistic values: *"For example, if a journalist... I would still read it if you know he looked it up himself, scanned it in himself for example. He just accelerated it, so to speak".*

Overall, these remarks reveal a shared ethical consideration: AI can be a useful tool, but human oversight is an important element of a functioning accountability mechanism. Whether through disclosure, oversight, or the regulation of liability, participants consistently highlighted the need to clearly define who is responsible when AI is used in journalistic contexts.

5.3 Concerns about AI in journalism

Where do these needs for rights and accountability come from? The individual statements made show that this could often be directly linked to societal concerns or concerns if no regulation occurs, that the participants held: Subject 4 (Group 2): *"Maybe also the pictures I post on social media. I don't want some form of AI to be able to do something with that, to use that or pluck that from the Internet and distribute that. So my face, which on my social media, that nothing should be done with that without my permission".* A second participant stressed: Subject 3 (Group 2): *"But it's a deepfake. Then it does get scary. Linked to that that I'm very afraid that eventually the population becomes very distrustful. So that now we can still pretty much all trust each other but at some point you can no longer tell whether*

something is real or fake. I am a little afraid that we will all be very distrustful of each other and of everything we send". Connected to this, the majority of individuals were worried about false and misleading information generated by AI and not being able to recognize it as such. For example, Subject 3 (Group 1): *"Well, not being able to recognize fake news. Not recognizing it yourself and making bad choices based on the information"*. Similarly, Subject 4 in Group 2: *"...unreliable source, that you therefore cannot distinguish whether something is real or fake"*. Together, these statements reveal that the need for regulation and accountability is deeply grounded in personal and societal fears. Individuals expressed fear about the loss of control over their digital identities, pointing to a fundamental concern about privacy. Beyond personal vulnerability, participants voiced a deeper collective concern: that AI technologies, like deepfakes, could erode trust. These insights suggest that the wish for regulation and responsibility is a response to a perceived crisis of trust, both in technology and in one another, highlighting a growing unease about the impact of AI if left unregulated.

6 Discussion

Even though citizens will be directly affected by the growing proliferation of generative AI in general, and journalism in specific, their voices are rarely heard when governance solutions are discussed. This qualitative study reveals a strong individual demand for regulation, accountability, and transparency in the use of AI in journalism (Dignum 2019). Participants expressed a clear desire for rights and regulatory frameworks to mitigate the risks of AI-generated content, particularly in ensuring the accuracy of news. Participants further emphasized the need for human oversight, institutional reliability, and mechanisms to hold individuals or institutions accountable when AI use leads to threats, such as misinformation. Thus, amidst calls for deregulation and an increasingly dominant narrative in European and international tech policy that regulation will have negative effects for innovation, and hence ultimately for citizens, this study paints a different picture. Drawing on the perspectives of participants in this study, innovation with AI in the media and some form of governance go hand in hand to secure the legitimate interests of citizens.

One key insight is that citizens do want to see regulation of the use of AI within journalism. This corroborates findings for the UK and Germany (König 2023; Modhvadia 2023). Confirming these findings, the results of this study add a comparative dimension, highlighting similarities in citizens' regulatory preferences across different countries.

While there was no clear preference for regulation by the government versus self-regulation, participants favored

strong oversight by news organizations themselves, guided by binding rules designed to protect the public from harms such as inaccurate information and manipulation.

In our study, Dutch individuals focused in particular on risks related to the lack of integrity and concerns around manipulation and false news, as well as the lack of human oversight. In terms of regulatory interventions, transparency was mentioned, and here, interestingly, the need for standardization and some official general labeling scheme was put forward. This finding is interesting insofar as, in practice, currently the different media organizations seem to handle each their own set of guidelines and labeling practices, and there is little industry standardization taking place. From the perspective of individuals, broader, industry-wide initiatives like the Nordic AI transparency initiative are welcome, but the desire for "same rules for everyone" goes beyond (Lagercrantz et al. 2024).

In terms of concrete transparency needs, the need for the disclosure of sources was highlighted. Such a disclosure of sources goes beyond the AI Act's approach to merely disclose whether or not AI has been used, and points to the need for additional information that helps individuals to decide whether or not to trust AI-generated content. This desire is in line with the arguments of Wittenberg et al. (2024), who stress that AI labels should inform individuals about the process by which news content was created rather than the simple fact that AI was involved. In a similar direction goes the suggestion by participants that a label could signal compliance with rules for the responsible use of AI. Again, pieces of information like these can help individuals to make a more advanced judgment than simply informing them that content has been AI generated.

Next to transparency, participants emphasized human oversight and verification as important measures to mitigate risks and ensure accountability. Interestingly, the previous research found that transparency and human oversight were also considered by media professionals as the most important measures to guarantee the responsible use of generative AI (Diakopoulos et al. 2024). Insofar, individuals seem to exhibit considerable trust in primarily journalists to be able to recognize and address problems, such as misleading or false content. Whether or not that trust is justified is another question that merits further research. The fact that participants also stressed the importance of institutional safeguards as well as wanting to have the ability to complain to someone if something was 'not correct', however, points to a shared responsibility between individual journalists but also the media organization. These findings relate to what has been described as the concept of the 'moral crumple zone' (Elish 2019), where responsibility in AI-systems is disproportionately placed on human beings despite their limited control over algorithmic decisions. Therefore, while calls for human oversight are important and legitimate, they must be

supported by careful consideration of how responsibility is distributed, and often distorted, within human–AI systems.

More research is necessary to further explore the specific mechanisms and frameworks that can effectively incorporate citizen voices into the governance of AI in journalism, given the strong preference for participatory processes highlighted by the participants of this study. Similarly, investigating institutional safeguards and complaint mechanisms could shed light on how shared responsibilities between journalists and media organizations can be structured to ensure responsible AI use. Finally, and interestingly, participants also discussed the potential of having some kind of registry or ‘court system’ for computers. These solutions demonstrate that involving citizens can also stimulate out-of-the-box thinking.

Even though the insights gained from this study are valuable, several limitations have to be acknowledged. First, the sample size was relatively small, consisting of 21 participants across three focus groups. While the diversity in age, gender, and education levels enhances the representativeness of the sample, the findings may not be generalizable to the broader Dutch population or citizens of other countries. Further quantitative studies with larger and more varied samples are needed. Second, the focus group methodology, while beneficial for observing interactions and eliciting detailed responses, may also introduce group dynamics that could influence individual opinions. Participants might conform to dominant views within the group, potentially skewing the data. To avoid this, future research should consider individual in-depth interviews, allowing the participants to express their views in a more private setting. Finally, the thematic analysis, although rigorous, is inherently subjective, and the coding process might reflect the researchers' biases despite efforts to ensure validity through consensus coding. To counteract this subjectivity, future work could combine these qualitative insights with quantitative survey data, aiming to verify the prevalence of regulatory preferences.

7 Conclusion

This study is one of the first to offer in-depth insights into individuals' expectations regarding the regulation of the use of AI in journalism. It reveals a strong demand for transparency, accountability, and human oversight. Individuals' concerns about misinformation, the erosion of trust and the need for standardized disclosures highlight the importance of aligning governance with the protection of citizens, whether it is through self-regulation or more formal regulation, like mandatory disclosure obligations for AI-generated information in the European AI Act. This study also demonstrates that there can be considerable discrepancies in what regulators or media organizations believe citizens should be informed about, and the kind of information citizens would

like to receive. For example, while the European AI Act and also several professional journalistic guidelines require labeling content as 'AI generated', participants expressed a need for more detailed information, including consistent source disclosure and transparent internal guidelines of how a media organization uses AI responsibly. Insofar, our findings can also inform the implementation of mandatory or self-regulatory transparency obligations. More generally, the study reveals further tensions in how citizens perceive institutional legitimacy, responsibility, and trust in the age of AI-driven journalism. While AI is often seen as a helpful tool, participants highlighted that its use in journalism must not only be governed by clear ethical guidelines, human oversight, and enforceable standards, but that the interests of citizens, too, should be more actively taken into account, making a plea for more participatory forms of governance.

8 Conflict of interest

The authors declare no competing interests.

Appendix A: Consensus codebook

Concerns of AI use & journalism

Informational concerns

Lack of accuracy and truthfulness (fake news)

Issues related to accuracy, truthfulness, and the potential for AI to spread fake news or polarize information.

Malintent/abuse

Concerns about AI being used with harmful intentions.

Polarization

Concerns about AI contributing to (societal) polarization.

Difficulties recognizing AI vs. non-AI information

Challenges in distinguishing between AI-generated and human-created content.

No or less diversity and aims for inclusivity in AI

No longer aiming for diversity or a myriad of viewpoints. Concerns of AI generating false information and not having the ability to verify the content.

Lack of authenticity in AI

Doubts about the authenticity of AI-generated content. Only code when mentioned explicitly.

Distrust AI

Distrusting content and potentially linked to the entity behind the content. Distrust triggered by AI content.

Existential fear of AI

Mainly on an individual level (losing your job).

No human oversight/losing control

Not being able to verify the context + fear of losing control.

Societal concerns

(Other) concerns regarding societal risks.

Other concerns

Additional concerns.

Opportunities of AI use & journalism

AI as a tool

Better able to sift through content/analyze

AI as being better for analysis, large chunks of content, pattern recognition.

AI as a fact checker

The role of AI in improving the accuracy of information.

Improving/adapting writing.

AI's potential capability to adapt—improve writing styles.

More diversity

AI as being better in generating diverse content, changing tone of voice, different perspectives etc.

More objectivity

AI's as a potential producer of more objective content.

Other opportunities.

Additional opportunities, e.g., automation.

General information needs

Transparency

Clear labeling

Importance of clearly marking AI-generated content.

Source reference

Desire for AI content to include source references and authorship information.

Verification, Regulation and Oversight

Human oversight

Need for AI content to be reviewed and (fact-)checked by humans.

Institutional reliability

A recognized label/logo connected to an institution.

Standards and guidelines

Establishment of guidelines for AI use in journalism to ensure credibility and accuracy.

Practical requirements for info needs

Logos and watermarks

Use of visible logos and watermarks to indicate AI-generated content.

Contrasting colors

Use of contrasting colors to make AI indicators stand out.

Placement of indicators/labels

Placement of AI indicators at the beginning of articles or prominently within the content.

Use differentiation

Different uses require different forms of disclosure. (e.g., grammar vs. generation of entire article)

Topic differentiation

Different topics require different forms of disclosure and different context.

News outlet differentiation

Different outlets require different forms of disclosure.

Ineffectiveness of labeling

Labeling can be inefficient.

Desired information specification

When reading an article, this information needs to be present: author, data used, articles used, origin of the content (legacy vs. other media), and level of AI

Journalism profession impact of AI

The impossibility of AI in news

AI cannot create news.

Job displacement

Concerns about AI replacing human journalists.

6. Regulatory entitlement (incl. Individual rights)

Accountability & responsibility (incl. punishment—enforcement)

Ensuring accountability and responsibility in AI use.

Protection against risks

Safeguards needed to protect against AI-related risks.

Other regulatory entitlement

Triggered behavior after disclosure

Skepticism

Consuming news more skeptically. With a grain of salt.

Distrust

Losing trust completely, stopping with reading. More extreme actions.

Individual fact checking (critical evaluation & validation)

Authenticity

When coding definition add who can display this authenticity (humans, journalists, AI).

Definition of Authenticity

Personality

AI's ability to exhibit personality.

Creativity

AI's capacity for creative output.

Factuality

AI's role in ensuring factual correctness.

Human contact

The importance of maintaining human interaction in content creation

Other.

Miscellaneous

Personal experiences with AI tools

Examples of AI usage, such as using ChatGPT for studying.

Double standards AI vs. non-AI

(e.g., fake news can also be created by humans).

Future developments of AI

Considerations on how AI will evolve.

Other anecdotes on AI.

Acknowledgements The authors would like to thank Sara Spaargaren from the AI, Media & Democracy Lab for her help facilitating the focus groups.

Author contributions All authors made substantial contributions to the conception and design of the work. Morosoli drafted and led the work and did the data interpretation. Data collection was done by Morosoli & Nautds. Morosoli, Nautds, Cools, and Venkatraj all collaborated for the data analysis. All authors reviewed the results and approved the final version of the manuscript.

Data availability No datasets were generated or analyzed during the current study.

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

- AlgoSoc. (2024). AlgoSoc AI opinion monitor. <https://monitor.algosoc.org/en/>
- De Andrade N, Galindo L, Zarra Heal J, Rom S (2023) Towards informed AI interactions: assessing the impact of notification styles on user awareness and trust. Open loop artificial intelligence act: a policy prototyping experiment operationalizing the requirements for AI systems—Part III. https://openloop.org/wp-content/uploads/2023/06/AI_Act_Towards_Informed_AI_Interactions.pdf
- Becker KB, Simon FM, Crum C (2025) Policies in parallel? A comparative study of journalistic AI policies in 52 global news organisations. Digit Journal. <https://doi.org/10.1080/21670811.2024.2431519>
- Beckett C (2023) Preparing for the coming wave of generative AI in journalism. LSE Blog. <https://blogs.lse.ac.uk/impactofsocialsciences/2023/09/18/preparing-for-the-coming-wave-of-generative-ai-in-journalism/>
- Braun, V., & Clarke, V. (2013). *Successful Qualitative Research a practical guide for beginners*. SAGE.
- Campbell JL, Quincy C, Osserman J, Pedersen OK (2013) Coding in-depth semistructured interviews: problems of unitization and intercoder reliability and agreement. Sociol Methods Res 42(3):294–320. <https://doi.org/10.1177/0049124113500475>
- Chen C, Gong X, Liu Z, Jiang, W, Goh SQ, Lam KY (2025) Trustworthy, responsible, and safe AI: a comprehensive architectural framework for ai safety with challenges and mitigations (No. arXiv:2408.12935). arXiv. <https://doi.org/10.48550/arXiv.2408.12935>
- Clarke V, Braun V (2017) Thematic analysis. J Posit Psychol 12(3):297–298. <https://doi.org/10.1080/17439760.2016.1262613>
- European Commission (2021) Netherlands AI strategy report. European Commission. https://ai-watch.ec.europa.eu/countries/netherlands/netherlands-ai-strategy-report_en
- Cools H, Diakopoulos N (2023) Towards guidelines for guidelines on the use of generative AI in newsrooms. Medium.Com. <https://generative-ai-newsroom.com/towards-guidelines-for-guidelines-on-the-use-of-generative-ai-in-newsrooms-55b0c2c1d960>
- Cools H, Diakopoulos N (2024) Uses of generative AI in the newsroom: mapping journalists' perceptions of perils and possibilities. J Pract. <https://doi.org/10.1080/17512786.2024.2394558>
- Cools H, Koliska M (2024) News automation and algorithmic transparency in the newsroom: the case of *the Washington Post*. Journalism Stud 25(6):662–680. <https://doi.org/10.1080/1461670X.2024.2326636>
- Department for the Coordination of Algorithmic oversight (DCA) (2024) AI & Algorithmic Risks Report Netherlands. Periodic insight into risks and effects of the use of AI and algorithms in the Netherlands.
- Diakopoulos N, Cools H, Li C, Helberger N, Kung E, Rinehart A, Gibbs L (2024) Generative AI in journalism the evolution of newswork and ethics in a generative information ecosystem. <https://doi.org/10.13140/RG.2.2.31540.05765>
- Diakopoulos N, Koliska M (2017) Algorithmic transparency in the news media. Dig J 5(7):809–828. <https://doi.org/10.1080/21670811.2016.1208053>
- Dignum V (2019) Responsible artificial intelligence: how to develop and use AI in a responsible way, vol 2156. Springer

- Dodds T, Vandendaele A, Simon FM, Helberger N, Resendez V, Yeung WN (2024) The impact of knowledge silos on responsible AI practices in journalism (No. arXiv:2410.01138). arXiv. <https://doi.org/10.48550/arXiv.2410.01138>
- Elish MC (2019) Moral crumple zones: cautionary tales in human-robot interaction. *Engag Sci Technol Soc* 5:40–60. <https://doi.org/10.17351/ests2019.260>
- Kaminski ME (2019) The right to explanation, explained. <https://doi.org/10.15779/Z38TD9N83H>
- Karlssoon M (2010) Rituals of transparency: evaluating online news outlets' uses of transparency rituals in the United States, United Kingdom and Sweden. *Journalism Stud* 11(4):535–545. <https://doi.org/10.1080/14616701003638400>
- Karlssoon M (2021) Transparency and journalism: a critical appraisal of a disruptive norm. Routledge
- König PD (2023) Sustainability challenges of artificial intelligence and citizens' regulatory preferences. *Gov Inf Q*. <https://doi.org/10.1016/j.giq.2023.101863>
- Lagercrantz A, Mattsson T, Neld K, Hansson V, Stjärne H, Benkö C, Ahlqvist C, Taubert J, Wennberg J, Savage J, Sandhammar K, Berg M, Careborg A, Karlsson-Bernfalk M, Enström A, Junkka Å, Elwin C, Edberg U, Österlund M, ... Rydgren Å (2024) Explorative industry collaboration on AI transparency in Swedish media. *Nordic AI Journalism*. <https://www.nordicaijournalism.com/ai-transparency>
- Lechterman TM (2022) The concept of accountability in AI ethics and governance. In: Bullock JB, Chen Y-C, Himmelreich J, Hudson VM, Korinek A, Young MM, Zhang B (eds) *The oxford handbook of AI governance*, 1st edn. Oxford University Press, pp 164–182
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2024). *Overheidsbrede visie Generatieve AI*. <https://www.rijksinnovatiecommunity.nl/actueel/nieuws/2024/01/19/visie-ai#:~:text=Het%20Nederlandse%20kabinet%20heeft%2018,disruptieve%20en%20tegelijk%20kansrijk%20technologie>
- Modhvadia R (2023). How do people feel about AI? A nationally representative survey of public attitudes to artificial intelligence in Britain. <https://www.adalovelaceinstitute.org/report/public-attitudes-ai/>
- Morgan DL (1996) Focus groups. *Annu Rev Sociol* 22(1):129–152
- Mulgan R (2000) 'Accountability': an ever-expanding concept? *Public Adm* 78(3):555–573
- Naudts, L., Helberger, N., Veale, M., & Sax, M. (2025). *A Right to Constructive Optimisation: A Public interest approach to recommender systems in the Digital Services Act*. https://doi.org/10.31235/osf.io/672hy_v1
- Naudts L, Dewitte P, Ausloos J (n.d.). Meaningful transparency through data rights: A multidimensional analysis
- Newman N, Fletcher R, Robertson CT, Ross Arguedas A, Nielsen RK (2024) Reuters Institute digital news report 2024. Reuters Inst Study J. <https://doi.org/10.60625/RISJ-VY6N-4V57>
- Novelli C, Taddeo M, Floridi L (2024) Accountability in artificial intelligence: what it is and how it works. *AI Soc* 39(4):1871–1882. <https://doi.org/10.1007/s00146-023-01635-y>
- OECD (2024) Exploring new frontiers of citizen participation in the policy cycle. Discussion paper. 2024 OECD global forum on building trust and reinforcing democracy
- Opgenhaffen M, d'Haenens L (2015) Managing social media use: whither social media guidelines in news organizations? *Int J Media Manage* 17(4):201–216. <https://doi.org/10.1080/14241277.2015.1107570>
- Paik S (2023) Journalism ethics for the algorithmic era. *Digit J*. <https://doi.org/10.1080/21670811.2023.2200195>
- European Parliament, the Council and Commission (2022) European declaration on digital rights and principles.
- Peek L, Fothergill A (2009) Using focus groups: lessons from studying daycare centers, 9/11, and Hurricane Katrina. *Qual Res* 9(1):31–59. <https://doi.org/10.1177/1468794108098029>
- Piasecki S, Morosoli S, Helberger N, Naudts L (2024) AI-generated journalism: Do the transparency provisions in the AI Act give news readers what they hope for? *Internet Policy Rev*. <https://doi.org/10.14763/2024.4.1810>
- Sadek M, Kallina E, Bohné T, Mougnot C, Calvo RA, Cave S (2025) Challenges of responsible AI in practice: scoping review and recommended actions. *AI Soc* 40(1):199–215. <https://doi.org/10.1007/s00146-024-01880-9>
- Simon FM (2024) Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena (Tow Report) [Columbia Journalism Review].
- Strycharz J, Ausloos J, Helberger N (2020) Data protection or data frustration? Individual perceptions and attitudes towards the GDPR. *Eur Data Prot Law Rev* 6(3):407–421. <https://doi.org/10.21552/edpl/2020/3/10>
- Sun M, Hu W, Wu Y (2022) Public perceptions and attitudes towards the application of artificial intelligence in journalism: from a China-based survey. *Journalism Pract*. <https://doi.org/10.1080/17512786.2022.2055621>
- Thynne I, Goldring J (1987) Accountability and control: Government officials and the exercise of power. (*No Title*)
- Toff B, Simon FM (2023) "Or they could just not use it?": The Paradox of AI Disclosure for Audience Trust in News. <https://doi.org/10.31235/osf.io/mdvak>
- Tyson A, Kikuchi E (2023) Growing public concern about the role of artificial intelligence in daily life. Pew Research Center. <https://www.pewresearch.org/short-reads/2023/08/28/growing-public-concern-about-the-role-of-artificial-intelligence-in-daily-life/>
- Varošane I (2022) On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI. *Int Rev Law Comput Technol* 36(2):95–117. <https://doi.org/10.1080/13600869.2022.2060471>
- Weymouth R (2015) Deliberative collaborative governance as a democratic reform to resolve wicked problems and improve trust
- Wittenberg C, Epstein Z, Berinsky AJ, Rand DG (2024) Labeling AI-Generated Content: Promises, Perils, and Future Directions. An MIT Exploration of Generative AI. <https://doi.org/10.21428/e4baedd9.0319e3a6>
- Wong J, Morgan D, Straub VJ, Hashem Y, Bright J (2022) Key challenges for the participatory governance of AI in public administration. <https://doi.org/10.31235/osf.io/pdcrm>

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