UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Understanding AI Disclosure Needs for News Production and Journalism

Venkatraj, Karthikeya Puttur; Morosoli, Sophie; Cools, Hannes; Naudts, Laurens; de Vreese, Claes; Helberger, Natali; Cesar, Pablo; El Ali, Abdallah

DOI
10.1145/3771882.3771899

Publication date
2025

Document Version
Final published version

Published in
Proceedings of MUM 2025

License
CC BY-NC-ND

Link to publication

Citation for published version (APA):
Venkatraj, K. P., Morosoli, S., Cools, H., Naudts, L., de Vreese, C., Helberger, N., Cesar, P., & El Ali, A. (2025). Understanding AI Disclosure Needs for News Production and Journalism. In S. Sorce, P. Elagroudy, & M. Khamis (Eds.), *Proceedings of MUM 2025: The 24th International Conference on Mobile and Ubiquitous Multimedia* (pp. 202-208). Association for Computing Machinery. https://doi.org/10.1145/3771882.3771899

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:10 Oct 2026

Understanding AI Disclosure Needs for News Production and Journalism

Karthikeya Puttur Venkatraj
Centrum Wiskunde & Informatica
Amsterdam, Netherlands
TU Delft
Delft, Netherlands
karthikeyapv16@gmail.com

Sophie Morosoli
University of Amsterdam
Amsterdam, Netherlands
s.v.morosoli@uva.nl

Hannes Cools
University of Amsterdam
Amsterdam, Netherlands
h.cools@uva.nl

Laurens Naudts
University of Amsterdam
Amsterdam, Netherlands
KU Leuven
Leuven, Belgium
l.p.a.Naudts@uva.nl

Claes de Vreese
University of Amsterdam
Amsterdam, Netherlands
C.H.deVreese@uva.nl

Natali Helberger
University of Amsterdam
Amsterdam, Netherlands
n.helberger@uva.nl

Pablo Cesar
Centrum Wiskunde & Informatica
Amsterdam, Netherlands
TU Delft
Delft, Netherlands
garcia@cwi.nl

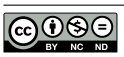
Abdallah El Ali
Centrum Wiskunde & Informatica
Amsterdam, Netherlands
Utrecht University
Utrecht, Netherlands
aea@cwi.nl

Abstract

Artificial Intelligence (AI) is revolutionizing the way content is produced and integrated into journalistic workflows. The EU AI act's Article 50 sets up transparency requirements aimed at encouraging the adoption and disclosure of AI in an ethical and responsible manner. In this study, we organized focus group interviews with Dutch citizens (N=21) to understand their expectations and needs regarding AI disclosures in the context of news production and journalism. These conversations are essential to understand if legal and regulatory policies are grounded in real-world experiences of citizens, and adequately address their concerns and enhance their digital interactions. We found that citizens predominantly favor disclosures of AI usage in journalistic content, in the form of (1) source references, (2) visual indicators (logos/watermarks) and (3) have varying preferences regarding information presentation and interaction modalities. Our findings highlight the need for interdisciplinary approaches to align standardization efforts with AI disclosures for news media.

CCS Concepts

• **Human-centered computing → Human computer interaction (HCI)**.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
MUM '25, Enna, Italy
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2015-4/25/12
https://doi.org/10.1145/3771882.3771899

Keywords

EU AI Act, generative artificial intelligence, AI disclosures, transparency, News, Journalism

ACM Reference Format:
Karthikeya Puttur Venkatraj, Sophie Morosoli, Hannes Cools, Laurens Naudts, Claes de Vreese, Natali Helberger, Pablo Cesar, and Abdallah El Ali. 2025. Understanding AI Disclosure Needs for News Production and Journalism. In *24th International Conference on Mobile and Ubiquitous Multimedia (MUM '25), December 01–04, 2025, Enna, Italy.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3771882.3771899

1 Introduction

The global technological revolution accelerated by Artificial Intelligence (AI) needs the pillar of citizen trust to safeguard against the consumption and spread of fake media. The World Economic Forum's 2024 Global Risks Report [1] ranked "misinformation and disinformation" as the highest global risk expected over these upcoming years. The European Union AI Act aims to create the conditions for responsible AI adoption through a risk-based regulatory approach [26]. Article 50 of the AI Act formulates a general commitment towards transparency about the use of AI in producing content, and stipulates under what conditions users need to be informed when media organizations use AI to produce audio, video, and under certain circumstances, also text. That being said, the AI Act fails to provide concrete guidance on how users should be informed and the types of disclosures that are conducive to the regulatory goal[13, 30], also considering the increasingly pervasive rise of the secretive use of large language models [44]. Disclosure guidelines are constantly evolving, with different flavors of visuals such as stars or sparkles icons that represent AI related features

[24], taxonomies such as 'Made with AI' [4] and 'Altered or Synthetic content'[17], and interactions, based on the organizations approach towards transparency. While they represent essential steps towards transparency, the lack of standardization will lead to a spur of different types of notifications and labels, similar to cookie consent banners on websites, that may end up confusing the user more [19, 27]. In addition, previous research demonstrated that users might find the use of AI to generate news without being transparent as misleading [15, 36]. Likewise, individuals seek the ability to filter based on the source, distinguishing between AI-generated and human journalist produced content [36].

Therefore it is essential to design AI disclosures in a way that provides meaningful information to users for their interaction with AI-generated content. We therefore ask: What are Dutch Citizens' information disclosure and transparency needs for AI-generated content in news media? To answer this, we conducted three focus group sessions with citizens (N = 21) of The Netherlands regarding their perspectives on the use of AI and its disclosure, specifically in the context of news media.

We found that: (1) participants want to be informed that they are consuming content generated by AI. (2) Participants want source references for the AI-generated text in the news article, so that they can verify its credibility. (3) Participants advocate for attention capturing disclosures such as using contrasting colors/sizes or pop-up windows. Lastly, participants indicate varying preferences in terms of disclosure types (text/icons/watermarks), placement (above/below articles), level of AI use (fully/ partially generated) and modality (text/audio/visual) for the disclosure, highlighting the need for disclosure personalization. These findings emphasize the need for interdisciplinary research to establish equilibrium between standardization and adaptability of AI disclosures within news media.

2 Related Work

2.1 Disclosure practices and strategies

When we assess trust towards the use of AI in journalism, audiences tend to avoid sharing news with headlines labeled as AI-generated [2], similarly Toff and Simon showed that news generated with the help of AI is perceived as less trustworthy, although this effect is counteracted when a list of sources used to generate the article is provided to the users [39]. Interestingly, we also see that journalistic content labeled as AI is perceived more trustworthy since AI is perceived to lack human biases and is hence more credible [21]. Evidently, these findings are yet inconclusive and need to be carefully assessed to unravel the complexity of using AI in news media.

Important contextual details regarding the origin of an artifact need to be provided to its users as an indicator of reliability [38]. To that end, certification labels are considered a robust method which is field tested in other domains such as the food and agriculture sector [18, 43], to communicate essential information to users using simple language, color codes and icons. However, extrapolation of this method to AI systems can foster 'Blind-trust' if correct performance indicators of the system are not implemented in the label [37]. Epstein et al. [14] highlight that terminology primarily depends on the goal of labeling the content. If the objective is to inform the user then terms such as 'AI Generated', 'Artificial' provide the best results. If it is to alert users about the content then terms such as 'Not real' and 'Manipulated' provide the most successful results. Interestingly, Chinese participants in this same study reported different interpretations of the term 'Artificial'. So it is imperative to ensure that the terminology used for labeling is consistent in interpretation within different cultural and language demographics [42]. The study by Epstein et al. [14] also revealed that the 'Implied Authenticity' effect might occur when labeling AI-generated content (cf., [35]), leading to any content that is not labeled to be considered as more authentic and not coming from AI.

To counter the proliferation of misinformation and distrust, Wittenberg et al. [42] put forward a framework for labeling AI-generated media based on a process based versus harm based transparency approach. Subsequent research using this framework showed that content creators were open to adopting such tools to share their creative processes using AI to their audiences [6]. Gils at al. [16] provide guidelines for designing labeling solutions for AI-generated content specifically in light of the EU AI Acts transparency provisions. They suggest designing disclosures that are accessible to a diverse audience, particularly considering individuals with disabilities. They encourage to use a layered approach to reveal the information to users, which helps avoid information overload (cf., [13]), while leveraging interactive elements to enhance the clarity of disclosure.

2.2 AI 'in' & 'for' News

AI's role in news content production has accelerated, owing to the boost in productivity and efficiency [32]. Journalists in newsrooms most commonly use AI for translations and information gathering [10, 33]. While this adoption is encouraged by technologically driven audiences, trust primarily drives the relationship that audiences have towards AI-generated content, for journalists and news outlets that use AI [31]. Stakeholders of the media industry advocate for responsibly developed ethical guidelines and its effective implementation. Additionally, appropriate training to personnel, curation of AI tools through reliable auditing and testing processes are necessary strategies to ensure responsible AI use [12]. A recent survey by Reuters found that majority of the public in six countries want some kind of disclosure practice to be adopted for AI use within new content production [15]. This is in strong contrast to the EU AI Act's provision (Article 50, para 5), that exempts news organizations from AI disclosure obligations after editorial review. Moreover current strategies to battle misinformation such as fact-checking, media literacy tips and news coverage of misinformation lead to increased skepticism towards all information online [20]. These sentiments underscore the dynamic trends of the technology, where the focus of the media industry should be to maintain and grow its relationship with audiences through open and trust-based communication. The first step of which is to understand what information the audience needs in this rapidly evolving landscape of AI integration within news content production.

3 Methods

3.1 Focus groups: approach and objectives

We conducted three focus group interviews in June 2024 to understand the information needs of Dutch citizens (N=21) and their views on transparency requirements for AI-generated content in the context of the media sector. We used the qualitative research approach of group interviews to create a setting to generate and discuss ideas using semi-structured questions to guide the discussion. Since focus groups are ideal for cases that require collecting diverse audience perspectives on a specific area of interest [34]. The interviews were structured into five parts where participant were provided the opportunity to deliberate regarding (1) concerns about using AI in news media, (2) characteristics of the information that needs to accompany AI-generated news, (3) behavioral traits shown when exposed to AI-generated content, (4) expectations towards AI-generated media compared to human generated media and, (5) Authenticity of AI-generated content. However, the sessions were not limited to these parts and were allowed to flow organically into other relevant and related topics. The predetermined interview guide is provided in supplementary material A. The focus groups were held on a university campus in Amsterdam. The group interviews were recorded, transcribed and translated from Dutch into English using a translation tool (DeepL [1]) after which a Dutch native speaker reviewed all the transcripts for accuracy.

3.2 Stimuli

To introduce parts 2-5 of the interviews, we created stimuli to contextualize possible scenarios of where and how the participants might engage with AI-generated news stories in the future. This was created by sourcing AI-generated images and text online from websites that were flagged as 'AI-generated news' by Newsguard [2] on topics of politics, sports and art. These topics were selected to take into account any domain specific news perceptions. Two types of news stimuli were created, (1) AI-generated headline along with AI-generated image and (2) AI-generated headline along with AI-generated text, since they are most commonly used formats to present short news stories online. We provided the participants with one stimulus of each type for each category of news topics, giving them a total of six AI-generated news stories. Photo editing software 'Photoshop' was utilized to match visual characteristics such as the placement of the headlines, images, text font size and social media icons to those found on news stories from legitimate websites, thereby ensuring realism to the content. Two stimuli, one on politics and one on sports, which were created and used in the sessions, are shown in Figure 1. All the stimuli used in during the group interviews are attached as supplementary material B.

3.3 Study procedure

The sessions began with an exercise where participants were asked by the interviewer to write down three concerns they had with respect to AI-generated content on a post-it note. They were prompted to share their concerns with the rest of the group and briefly elaborate on their views. Afterward, they proceeded to rank the post-its as a group until they reached a consensus on the order from most to least concerning. This exercise lasted for approximately 15 minutes after which the interviewer introduced the plenary discussion section along with the AI-generated news stimuli. Each session lasted for approximately 60 minutes. Throughout the discussion, the interviewer would intervene with relevant questions from the interview guide or clarifications to steer the discussion when required. The study protocol was approved by the University of Amsterdam's ethics review board.

3.3.1 Participants. 21 participants were recruited through a polling company and split into three groups of seven participants ensuring diversity of age, gender and education level in each of the focus groups. Knowledge about AI systems or exposure to AI-generated content was not a prerequisite for recruitment, since we wanted to capture the everyday citizens opinions about the topic. All participants were citizens of The Netherlands and were proficient in the Dutch language, in which the interviews were held. Seven were in the 23-26 age group, seven were in the 32-46 age group, seven were in the 51-63 age group. Dutch citizens have minimal trust in Automatic Decision-Making (ADM) systems to determine accurately the news one sees [9]. Additionally, previous work by Piasecki et al [36] showed that dutch citizens are open to using tools that aid in navigating news consumption. For these reasons, the Dutch population was selected for this study.

3.4 Analysis

We adopted a group based thematic analysis to identify relevant themes from the three focus groups [5]. The analysis was conducted by four coders (first four authors). First, the coders immersed themselves in the transcriptions and created four separate personal codebooks using open coding [3]. Using these as references, the coders then compiled a consensus codebook through deliberation, which would be used for the coding process. The consensus codebook consisted of nine main codes and 34 sub-codes, the full codebook has been provided in supplementary material C. Multiple coding sessions were held where the coders sat together and performed line-by-line analysis, discussing and assigning appropriate codes to every statement in the three focus group interview transcripts using the software NVivo. Codes were assigned through a "negotiated agreement" process, in which all coders discussed and deliberated on each assigned code and theme to ensure validity and inter-coder reliability [8]. Quotes used in this study are linked to participants by indicating which group (G1/G2/G3) they belonged to, followed by the specific participant number (P1-P7). Group and participant numbers are specified to illustrate the recurring themes and similarities in discussions across the different group sessions.

4 Findings

4.1 General information needs

Participants emphasized their need for clear labeling of AI-generated content, and their desire to know where the AI is gathering information from, for instance [G2-P6] mentioned *"I would think it is necessary that it (source of information) is actually there because then you just indeed know where it (news article) comes from."* and [G3-P5] suggests *"But what if it (AI) just put down a source? If it*

[1] https://www.deepl.com/en/translator
[2] https://www.newsguardtech.com/special-reports/newsbots-ai-generated-news-websites-proliferating/

(a) Political news article with AI-generated image and headline

(b) Sport news article with AI-generated text and headline

Figure 1: AI-generated news articles created as stimuli for the Focus group sessions

was indicated that it was generated with AI, based on this article, these posts, and then it's kind of a summary of five or six posts.". Participants also noted that sources are important to establish the relevance of the news article, *"Maybe also where they get it (news) from then. Of course, it could be a combination of articles, but whether something is really news, or whether it is rewritten old news."*[G1-P1]. Participants showed basic (non-technical) understanding of how AI works with data and various sources, for instance [G1-P2] tries to breakdown the complexity of source references stating that *"AI... that kind of text that's just, that could also sometimes be probably 1000 sources. That AI just picks words from articles that could just be 1000 articles. So you can never really cite the source. You can never have a source reference because so many sources. That's where I think if you really want to know if it's right, based on facts or if it's based on AI."*. There was consensus in the three groups that human oversight over AI systems is paramount to ensure accountability for news, *"That AI is controlled by a person. Not 100 percent of generative AI but a person should control for news."*[G3-P7]. This control should be augmented with review procedures that should take place for the AI-generated output, *"..you always have to do some kind of check and not just blindly throwing in a source. And think here: publish, enter, done."*[G1-P5]. Participants referred to (independent) 'institutions' that should be responsible for providing a 'seal of approval' for the AI system to be used for generating content that can be used for the media sector. For instance, a participant remarked that *"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, even fake logos in circulation and people are being misled anyway. And it has to be something, something official."*[G1-P6]. This discussion led to the need for the establishment of standards and guidelines to maintain the credibility of the news content being generated, [G1-P3] expressed that *"there is a kind of code of conduct for journalism, so then that should also be part of that, it's just the same rules for everyone and, if you use AI in journalism, 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."*.

4.2 Practical requirements for disclosure

Some participants expressed that the disclosure should take the form of an icon, for instance [G2-P6] mentioned *"As long as it is mentioned somewhere and that can be through an icon, that can be through text, but plain and simple, it is clear that it is indeed AI-generated"*. Other participants showed preference towards watermarks, for example [G3-P3] remarks that *"For me then, the preference would be for a watermark which is not removable, and also indeed, at the time you print it out still visible."* and other participants expressed that watermarks would indeed work better for specific cases such as AI-generated movies and images, [G2-P2] remarked *"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? A beautiful image of nature, landscape...and someone thinks, yeah, I'm not going to tell you that's AI-generated"*. Interestingly, in contrast to this sentiment participants also noted that visual indicators are ineffective for users who want to seek more information regarding the AI system that was used to generate the content, *"If you do know to what extent AI has been used then you do actually need to mention more than just an icon."*[G2-P7]. Participants discussed different ways to make the disclosure stand out to inform the users, [G1-P5] suggests using contrasting colors, *"So really something contrasting in terms of color, that it stands out and doesn't fall away in an image.."*. Some participants indicated that the disclosure should be in the form of a pop-up, *"I would maybe just as with if you have to accept cookies for a website, that you also just get a notification or some sort of Pop-up of: This was written by AI."*[G3-P2]. Others had specific preferences for the placement of the disclosure, [G3-P5] noted that *"If it's above the pictures, that still draws attention faster than between a patch of text. If you put it like that here for example and then it stands out even less than maybe at the bottom"*.

Similar to these aesthetic preferences, participants also commented on the differences in news topics and use cases which would require disclosure obligations. With respect to news topics, participants generally agreed that AI should not be used for political news, while it was considered favorable for data driven

news for shopping or sports. [G1-P7] commented that *"Say, one of those volleyball articles, yeah, that's a bit informative about how popular the sport is. Then I would think, I might believe AI. Because it probably has all the data on that over the past few years. But with something like that with Zelensky which is so timely and still so actively in the news every day, I would assume that less."* and [G1-P5] emphasized *"Because it can spread a lot of misinformation, which in turn can be dangerous if you're talking about, say, Zelensky. With volleyballs, I think it might be less unfortunate, but I can still imagine if someone buys that volleyball…"*. Participants highlighted the need for disclosure for certain unique cases, for instance [G3-P5] expressed *"Also depends on the context. You can also drop an already existing article into AI of course and say makes this bigger, make it a bit more comprehensive or translate from French to Dutch. Do you necessarily have to add a notification if it's just translated by AI?"*, to this [G3-P4] states *"Definitely (needed) with translations."*. In the first session [G1-P4] stressed that *"I think the most sensitive really are those photos. Imagine: They make an article about you, you're suddenly famous and then suddenly you're there in a photo and that's not you. That, I think that's very helpful for the person being written about that it then says 'AI', that everyone knows it's not real."* Regarding textual disclosures, participants were using different terms that should be used in the disclosure in all the sessions, [G1-P5] stated *"A source reference does seem useful. Normally you have something like an author, date and place. And then it should say, something like, 'generated by AI'."*, while [G1-P1] suggests *"..or in in indeed as we said, both in the beginning 'AI-generated', and also somewhere really a kind of label."* and [G3-P4] mentioned *"...Because a lot of times you don't look at who wrote it. And so if you do a little caption 'created by AI'..."*. Furthermore, [G2-P2] advocated for different levels of AI use to be mentioned within the disclosure, *"Maybe more options, in the sense of 'fully generated', 'partially generated', 'checked by AI'..."*.

5 Discussion

To enhance clarity and guide the reader, key design Considerations (C) are highlighted in bold and explicitly labeled in this section **(e.g., C.1, C.2, C.3)** to mark takeaways for easy reference.

5.1 Towards user-centric AI disclosures in journalistic contexts

This work aimed at understanding Dutch citizens' information needs regarding their exposures and interaction with AI-generated content. We extracted insights that provides nuance to existing body of research through focus group interviews with a representative sample of Dutch citizens. In these sessions most participants shared general concerns that they have concerning AI usage within journalistic contexts, such as the need for human oversight and reviews. Furthermore, **(C.1) participants collectively agree on the necessity of AI disclosures in news**. Research indicates that AI acceptance within the media context is driven by the audiences' perception of the technology and its applications [31]. Therefore, the audience first needs to be made aware that they are interacting with AI-generated content in order to then establish a reference point for its credibility within a domain as sensitive as news. The findings of the focus groups also showed that **(C.2) source references present a crucial method to strengthen the acceptability of AI-generated content**, which is in line with previous work by Epstein et al. [14]. Methods such as 'Origin tracing' can be leveraged to detect the core ideas that were used to derive the output of generative AI systems [29]. **(C.3) Some participants also highlighted the need for independent institutions that need to be setup for the standardization of disclosure practices**. They refer to **(C.4) 'seal of approval' that is given to systems or processes that follow a protocol ensuring reliable outputs**. This however may necessitate measures to encode provenance cryptographically, for example in images (cf., Coalition for Content Provenance and Authenticity [7]) or audio (SynthID [11]). Furthermore, we found that **(C.5) most participants had varied preferences regarding the potential presentation (icons / watermarks) of AI disclosures in news**. Nonetheless, **(C.6) all participants specified the need for clear labeling of AI-generated content**, signifying the need for a user-centric approach towards disclosure design. The key to effective AI disclosures might lie in user-controlled interface integration, moving from provider-constrained environments to user controlled designs. **(C.7) The disclosures should also be tuned to topic sensitivity, ensuring users understand both the presence and absence of AI**. However, determining sensitivity levels in real time can be challenging to implement. This type of disclosure would need specific technological capabilities such as: content flagging methods that work even when authors have not labeled AI content [41], thus circumventing the "implied truth effect" [14], accessible interface design [25], and overlay functionality such as pop-ups to actively inform users about AI detection. The strength of this type of approach is the power is placed in the users' hands to dictate their experience of transparency, while maintaining consistency across different content types. The usability of this approach however needs evaluation, since it poses potential risk to UX standardization efforts across the sector.

5.2 Study limitations

This study was conducted only within a Dutch context. While this provides valuable insights, it is limited to the media sector in The Netherlands and no comparison was made to other countries. Additionally participants expressed differing attitudes toward text and visual modalities, the stimuli created for the study did not differentiate between these formats. As a result, our findings do not reflect distinctions between modalities. Future studies using modality-specific stimuli are necessary to investigate these participant-reported differences. Furthermore, the context of this study was limited to the use of AI in journalism. Therefore, the results cannot be transferred to other uses of AI and related disclosures. Moreover, modern news media include more than just static text and images; formats such as videos, interactive graphics and audio are also common and can be presented in a variety of ways. Our study stimuli did not fully represent this diversity, which limits the generalizability of our findings. Lastly, these needs were identified using in a focus group setting, and may differ when people interact with content in situ on their own, posing a risk to its ecological validity. Additionally, initial questions about concerns regarding AI in news might have had a negative suggestion bias during the focus groups. Nevertheless, we believe our findings provide an important first step towards sense making in the information disclosure space.

5.3 Future Work

To ensure a future where AI disclosures are meaningful and effective within and beyond the media sector, we first need to understand and unpack the full UX design space of AI system disclosures, across modalities and domains (cf. [13]). As such, we envision that the designs themselves should be an outcome of iterative participatory AI [23, 28, 45] as well co-design sessions with industry stakeholders, including media organizations, citizens, and policymakers. Furthermore, while our work considered (generative) AI output, disclosures may extend well beyond static one-shot interactions and towards dynamic human-AI collaborations where it may not always be clear what output was produced by humans versus AI (cf., [22]), and how their respective configurations look like [40]. These are necessary steps to generating solutions grounded in real world scenarios that can bridge gaps between the needs of audiences and industry stakeholders while maintaining core journalistic and democratic values as generative AI content floods every media channel.

6 Conclusions

We conducted focus group interviews with Dutch citizens (N=21) to gather in-depth insights into their needs and expectations regarding disclosures of AI use for news production and journalism. This study was set up in the context of the EU AI act's transparency requirements of ethical and responsible disclosures of AI use. The findings of this study are important to highlight discrepancies between expectations of citizens and what news organizations or regulatory agencies believe citizens would want. Our findings show a clear and strong desire of transparency in AI-generated news content in the form of source references, that can validate the information. Participants also emphasized the need for industry-wide standardized disclosure practices which clearly communicate the use of AI using visual indicators (logos/watermarks). Finally, our findings highlight the need for interdisciplinary approaches to align standardization efforts with varying participant preferences towards information presentation and interaction modalities.

Acknowledgments

We thank Sara Spaargaren from AI, Media & Democracy Lab for her help in facilitating the focus groups. This publication is part of the AI, Media & Democracy Lab (Dutch Research Council project number: NWA.1332.20.009).

References

[1] World Economic Forum 2024. 2024. The Global Risks Report 2024. https://www.weforum.org/publications/global-risks-report-2024. [2024-1-19].
[2] Sacha Altay and Fabrizio Gilardi. 2024. People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. *PNAS Nexus* 3, 10 (10 2024), pgae403. arXiv:https://academic.oup.com/pnasnexus/article-pdf/3/10/pgae403/59961427/pgae403.pdf doi:10.1093/pnasnexus/pgae403
[3] Lucia Benaquisto. 2008. Open Coding. In *The SAGE Encyclopedia of Qualitative Research Methods*. SAGE Publications, Inc., Thousand Oaks. doi:10.4135/9781412963909.n299
[4] Monika Bickert. 2024. https://about.fb.com/news/2024/04/metas-approach-to-labeling-ai-generated-content-and-manipulated-media/
[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa doi:10.1191/1478088706qp063oa
[6] Olivia Burrus, Amanda Curtis, and Laura Herman. 2024. Unmasking AI: Informing Authenticity Decisions by Labeling AI-Generated Content. *Interactions* 31, 4 (June 2024), 38–42. doi:10.1145/3665321
[7] C2PA. 2020. Coalition for Content Provenance and Authenticity. https://c2pa.org/
[8] John L. Campbell, Charles Quincy, Jordan Osserman, and Ove K. Pedersen. 2013. Coding In-depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement. *Sociological Methods & Research* 42, 3 (2013), 294–320. doi:10.1177/0049124113500475
[9] AlgoSoc Consortium. 2025. *AlgoSoc AI Opinion Monitor*. AlgoSoc Consortium. https://monitor.algosoc.org/en/
[10] Hannes Cools and Nicholas Diakopoulos. 2024. Uses of Generative AI in the Newsroom: Mapping Journalists' Perceptions of Perils and Possibilities. *Journalism Practice* 0, 0 (2024), 1–19. arXiv:https://doi.org/10.1080/17512786.2024.2394558 doi:10.1080/17512786.2024.2394558
[11] Google DeepMind. 2024. Watermarking AI-generated text and video with SynthID. https://deepmind.google/discover/blog/watermarking-ai-generated-text-and-video-with-synthid/
[12] Nicholas Diakopoulos, Hannes Cools, N Helberger, C Li, Ernest Kung, Aimee Rinehart, et al. 2024. Generative AI in Journalism: The Evolution of Newswork and Ethics in a Generative Information Ecosystem.
[13] Abdallah El Ali, Karthikeya Puttur Venkatraj, Sophie Morosoli, Laurens Naudts, Natali Helberger, and Pablo Cesar. 2024. Transparent AI Disclosure Obligations: Who, What, When, Where, Why, How. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 342, 11 pages. doi:10.1145/3613905.3650750
[14] Ziv Epstein, Mengying Cathy Fang, Antonio Alonso Arechar, and David Gertler Rand. 2023. What label should be applied to content produced by generative AI? doi:10.31234/osf.io/v4mfz
[15] Richard Fletcher and R Nielsen. 2024. What does the public in six countries think of generative AI in news?
[16] Thomas Gils, Frederic Heymans, Wannes Ooms, and Jan De Bruyne. 2024. From Policy to Practice: Prototyping The EU AI Act's Transparency Requirements.
[17] Google. 2024. Google. https://support.google.com/youtube/answer/14328491?hl=en%5C&co=GENIE.Platform%5C%3DAndroid
[18] Klaus G Grunert, Sophie Hieke, and Josephine Wills. 2014. Sustainability labels on food products: Consumer motivation, understanding and use. *Food policy* 44 (2014), 177–189. https://doi.org/10.1016/j.foodpol.2013.12.001
[19] Hana Habib, Megan Li, Ellie Young, and Lorrie Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 621, 27 pages. doi:10.1145/3491102.3501985
[20] Emma Hoes, Brian Aitken, Jingwen Zhang, Tomasz Gackowski, and Magdalena Wojcieszak. 2023. Prominent Misinformation Interventions Reduce Misperceptions but Increase Skepticism. doi:10.31234/osf.io/zmpdu
[21] Lennart Hofeditz, Milad Mirbabaie, Jasmin Holstein, and Stefan Stieglitz. 2021. Do You Trust an AI-journalist? A Credibility Analysis of News Content with AI-Authorship.. In *ECIS*. Association for Information Systems, Marrakech, Morocco, 50. https://aisel.aisnet.org/ecis2021_rp/50
[22] Angel Hsing-Chi Hwang, Q. Vera Liao, Su Lin Blodgett, Alexandra Olteanu, and Adam Trischler. 2024. "It was 80% me, 20% AI": Seeking Authenticity in Co-Writing with Large Language Models. arXiv:2411.13032 [cs.HC] https://arxiv.org/abs/2411.13032
[23] Nanna Inie, Jeanette Falk, and Steve Tanimoto. 2023. Designing Participatory AI: Creative Professionals' Worries and Expectations about Generative AI. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 82, 8 pages. doi:10.1145/3544549.3585657
[24] Kate Kaplan. 2024. The proliferation and problem of the sparkles Icon. https://www.nngroup.com/articles/ai-sparkles-icon-problem/
[25] Nam Wook Kim, Shakila Cherise Joyner, Amalia Riegelhuth, and Yeeun Kim. 2021. Accessible visualization: Design space, opportunities, and challenges. *Computer graphics forum* 40, 3 (2021), 173–188.
[26] Johann Laux, Sandra Wachter, and Brent Mittelstadt. 2024. Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk. *Regulation & Governance* 18, 1 (2024), 3–32. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/rego.12512 doi:10.1111/rego.12512
[27] Adam Lee. 2021. The confusing world of cookie consent. https://www.and.digital/spotlight/blog/the-confusing-world-of-cookie-consent
[28] Jie Li, Hancheng Cao, Laura Lin, Youyang Hou, Ruihao Zhu, and Abdallah El Ali. 2024. User Experience Design Professionals' Perceptions of Generative Artificial Intelligence. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 381, 18 pages. doi:10.1145/3613904.3642114

[29] Linyang Li, Pengyu Wang, Ke Ren, Tianxiang Sun, and Xipeng Qiu. 2023. Origin tracing and detecting of llms.

[30] Zihao Li. 2023. Why the European AI Act transparency obligation is insufficient. *Nature Machine Intelligence* 5, 6 (2023), 559–560.

[31] Sophie Morosoli, Valeria Resendez, Laurens Naudts, Natali Helberger, and Claes de Vreese. 2025. "I Resist". A Study of Individual Attitudes Towards Generative AI in Journalism and Acts of Resistance, Risk Perceptions, Trust and Credibility. *Digital Journalism* 0, 0 (2025), 1–20. arXiv:https://doi.org/10.1080/21670811.2024.2435579 doi:10.1080/21670811.2024.2435579

[32] Amaya Noain-Sánchez. 2022. Addressing the Impact of Artificial Intelligence on Journalism: The perception of experts, journalists and academics. Communication and Society, 35 (3), 105–121.

[33] Andreas L Opdahl, Bjørnar Tessem, Duc-Tien Dang-Nguyen, Enrico Motta, Vinay Setty, Eivind Throndsen, Are Tverberg, and Christoph Trattner. 2023. Trustworthy journalism through AI. *Data & Knowledge Engineering* 146 (2023), 102182. doi:10.1016/j.datak.2023.102182

[34] Lori Peek and Alice Fothergill. 2009. Using focus groups: lessons from studying daycare centers, 9/11, and Hurricane Katrina. *Qualitative Research* 9, 1 (2009), 31–59. arXiv:https://doi.org/10.1177/1468794108098029 doi:10.1177/1468794108098029

[35] Gordon Pennycook, Adam Bear, Evan T. Collins, and David G. Rand. 2020. The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings. *Management Science* 66, 11 (2020), 4944–4957. arXiv:https://doi.org/10.1287/mnsc.2019.3478 doi:10.1287/mnsc.2019.3478

[36] Stanislaw Piasecki, Sophie Morosoli, Natali Helberger, and Laurens Naudts. 2024. AI-generated journalism: Do the transparency provisions in the AI Act give news readers what they hope for? *Internet Policy Review* 13, 4 (2024), 1–28.

[37] Nicolas Scharowski, Michaela Benk, Swen J. Kühne, Léane Wettstein, and Florian Brühlmann. 2023. Certification Labels for Trustworthy AI: Insights From an Empirical Mixed-Method Study. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 248–260. doi:10.1145/3593013.3593994

[38] Imani N. Sherman, Elissa M. Redmiles, and Jack W. Stokes. 2020. Designing Indicators to Combat Fake Media. arXiv:2010.00544 [cs.HC] https://arxiv.org/abs/2010.00544

[39] Benjamin Toff and Felix M. Simon. 0. "Or They Could Just Not Use It?": The Dilemma of AI Disclosure for Audience Trust in News. *The International Journal of Press/Politics* 0, 0 (0), 19401612241308697. arXiv:https://doi.org/10.1177/19401612241308697 doi:10.1177/19401612241308697

[40] Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. 2024. When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour* 8 (10 2024), 2293–2303. doi:10.1038/s41562-024-02024-1

[41] Debora Weber-Wulff, Alla Anohina-Naumeca, Sonja Bjelobaba, Tomáš Foltýnek, Jean Guerrero-Dib, Olumide Popoola, Petr Šigut, and Lorna Waddington. 2023. Testing of detection tools for AI-generated text. *International Journal for Educational Integrity* 19, 1 (2023), 26.

[42] Chloe Wittenberg, Ziv Epstein, Adam J. Berinsky, and David G. Rand. 2024. Labeling AI-Generated Content: Promises, Perils, and Future Directions. *An MIT Exploration of Generative AI* -, - (mar 27 2024). https://mit-genai.pubpub.org/pub/hu71se89.

[43] Wen Wu, Airong Zhang, Rieks Dekker van Klinken, Peggy Schrobback, and Jane Marie Muller. 2021. Consumer trust in food and the food system: a critical review. *Foods* 10, 10 (2021), 2490. https://pmc.ncbi.nlm.nih.gov/articles/PMC8536093/

[44] Zhiping Zhang, Chenxinran Shen, Bingsheng Yao, Dakuo Wang, and Tianshi Li. 2024. Secret Use of Large Language Model (LLM). arXiv:2409.19450 [cs.HC] https://arxiv.org/abs/2409.19450

[45] Douglas Zytko, Pamela J. Wisniewski, Shion Guha, Eric P. S. Baumer, and Min Kyung Lee. 2022. Participatory Design of AI Systems: Opportunities and Challenges Across Diverse Users, Relationships, and Application Domains. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 154, 4 pages. doi:10.1145/3491101.3516506