Tricky Trade-Offs on a Transparency Spectrum: How the *Financial Times* Approaches Transparency about AI Use in News

PRE-PRINT – NOT PEER-REVIEWED

Liz Lohn[1], Felix M. Simon[2]

[1]The Financial Times, London, United Kingdom, liz.lohn@ft.com

[2]Reuters Institute for the Study of Journalism, Oxford Internet Institute, Corpus Christi College, University of Oxford, Oxford, United Kingdom, felix.simon@politics.ox.ac.uk

7th November 2025

Abstract

As news organisations adopt artificial intelligence (AI), they face growing pressure to be transparent about when and how it is used – yet practical approaches remain uneven. This paper examines AI transparency through an in-depth case study of the *Financial Times* (FT). Drawing on 13 semi-structured interviews with 12 senior managers across editorial, product, data science, and communications, together with internal documents, we show that the FT approaches transparency as a hybrid of policy, process, and practice, framed by a desire to safeguard both internal and external trust. Transparency is calibrated to context: internally, AI use is signposted in tools and reinforced through training and personal accountability; externally, the prominence of disclosure scales with system autonomy and editorial oversight, with stronger labelling for no-human-in-the-loop features than for AI-assisted, journalist-edited outputs. We identify nine factors that shape audience-facing disclosure – legal/provider requirements, industry benchmarking, nature of the task, human oversight, system novelty, audience expectations, perceived risk, commercial sensitivities, and design constraints – and five crosscutting challenges, including site-wide consistency (especially on mobile) and potential "transparency backfire." Conceptually, our analysis links AI transparency to isomorphic pressures and to intersecting institutional logics. We argue that AI transparency in news is best understood as a spectrum, evolving with technological advancements, commercial, professional and ethical considerations and shifting audience attitudes.

Keywords: Artificial Intelligence, AI, Generative AI, Journalism, News, Trust, Transparency

Introduction

The future of professional news media is unstable. Public trust in journalism has been declining for years, and news avoidance is on the rise, raising concerns about the role of credible journalism in democratic societies (Communications and Digital Committee, 2024). At the same time, recent advances in artificial intelligence (AI) and its wider use in news have raised concerns that its use could further destabilise the already brittle trust in news.

Much of the public in various countries is sceptical about the use of AI in news (Newman et al., 2025, Simon et al., 2025) and about the technology more broadly. Some of this scepticism is justified. AI systems can discriminate, perpetuate biases or produce hallucinations. Various news outlets have used AI poorly and provoked public backlash. All this has intensified calls to establish robust norms and standards for transparency and accountability around the use of AI. Transparency is a critical ethical issue in journalism because many quality news organisations historically have acted as intermediaries of public knowledge and checks on power. However, to be effective, audiences must be able to trust not only the journalistic content itself but also the processes by which news is produced. Meanwhile, this trust is paramount for publishers from a business perspective, too.

Yet, despite the increasing use of AI in news, transparency practices around it remain inconsistent: while some news organisations openly disclose the role of AI in content production, others provide little to no indication of their AI use (Becker et al., 2025). The lack of standardised transparency procedures also raises questions about who is responsible for AI-use, accountable in the case of errors, and whether audiences can accurately assess the credibility of AI-assisted journalism. Additionally, within newsrooms, AI adoption may affect internal transparency – how journalists and editors are able to understand and monitor AI systems used to make and distribute the news.

All this creates urgency for studying approaches to AI transparency in news. This article offers an in-depth case study of how a leading news organisation – the *Financial Times* (FT) – approaches AI transparency. We explicate the motivations for AI transparency at the FT, the factors shaping both internal and external transparency around AI, as well as the challenges around this issue. Through interviews with 12 senior managers and experts from various departments and an analysis of internal documents we find that the *Financial Times* treats AI transparency as both an ethical imperative and a business necessity, operationalising it through a hybrid of policy, process, and practice underwritten by senior-leadership endorsement and a risk-sensitive, experiment-and-iterate approach. Transparency at the FT is seen as instrumental to safeguarding trust internally – via signposted tools, training, and communication to reduce frictions and surface errors early – and externally – through disclosures calibrated to nine factors shaping what, when, and how the FT discloses AI use: legal and provider requirements; industry benchmarking; degree of human oversight and AI contribution; the nature of the task; system novelty and uncertainty; audience expectations and research; perceived risk of harm or error; commercial sensitivities; and design constraints. These drivers sit alongside five cross-cutting challenges: achieving consistent, site-wide labelling (especially on mobile); overcoming silos and communication fatigue; keeping pace with evolving models and norms; guarding against human over-reliance; and mitigating audience misinterpretations or transparency backfire. Taken

together, our evidence adds granularity to prior research and shows that AI transparency is a lived, spectrum-based governance practice shaped by isomorphic pressures and intersecting professional, managerial, and commercial logics, rather than a binary or one-off decision.

Literature Review

AI in news

For the purpose of this paper, we define AI following the EU AI Act: "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" (Official Journal of the European Union, OJ L, 12 July 2024).

The use of AI in journalism dates back more than a decade (see e.g. Diakopolous, 2019) and can now be observed throughout the whole editorial process, across four phases: (1) news gathering (e.g. automated content aggregation), (2) news production (e.g. auto summarisation or voice synthesis), (3) news verification (checking the information), and (4) news distribution and moderation (e.g. content personalisation or search engine optimisation) (Simon, 2024 & 2025; Cools and Diakopoulos, 2024). Recent developments in the field of foundation models and LLMs, often referred to as Generative AI (GenAI), have further accelerated experimentation with AI in news, given these systems' given these systems' general-purpose capacity to generate and transform language, images, audio, and code; reason over multimodal inputs; and interact via natural language across tasks and domains (Bommasani et al., 2021; OpenAI, 2023). A growing number of publishers have sought to integrate such AI systems, given their ability to (assist in) draft(ing) and illustrate news stories, headlines; ingest, structure, and interrogate vast multimodal datasets to surface news-worthy patterns and facts; condense complex documents or live event streams into concise, accurate summaries on demand; translate, fact-check, and personalise content for diverse audiences); and integrate into end-to-end newsroom pipelines that support story ideation, production, and distribution at scale, to provide just some examples.

However, employing AI both poses risks and offers opportunities to news media organisations. Among the possibilities, there are (1) improvements to efficiency, (2) the scalability of work processes, and (3) better news personalisation (Simon, 2025) There are also risks: (1) the potential oversimplification of complex issues; (2) hallucinations and bias in outputs; (3) losing control of journalistic autonomy (Cools and Diakopoulos, 2024; Simon, 2022), as well as broader concerns such as the impact on the environment, growing dependency on the technology sector, and a restructuring of way citizens access and consume news and information as AI is integrated more deeply into digital media.

Additionally, there is a growing body of research showing that the public is sceptical about AI use in news in general (CNTI Briefing 1, 2025), and not comfortable with fully AI-generated news (Newman et al., 2025), while holding complex attitudes regarding AI use in news more generally (Mitova et al., 2025, Simon et al., 2025). With trust in news already low in many countries (Newman et al., 2025) and declining overall in recent years (Fletcher et al., 2025), fears have grown that the use of AI in news may further undermine trust in news, given audiences

complex attitudes to the topic. Consequently, publishers have been trying to find ways to mitigate these negative effects, including by being transparent about their AI use.

Transparency and trust in news

Transparency, as a foundational ethical norm in modern journalism (Deuze, 2005), relates to the relationship between journalists and the public, signalling to the latter that they can trust the news to present the truth. Transparency empowers audiences with means to better verify facts, confirm the credibility of news reporting, and hold journalism to account (Curry and Stroud, 2019; Karlsson, 2020). As such, it serves as one pre-condition for trust in news (Fawzi et al., 2021; Newman et al., 2025), which is critical in two ways: First, because trust (at least in an ideal scenario) enables news to fulfil its democratic function of informing people about matters of common concern, enabling deliberation, and holding power to account (compare Nielsen, 2017). Second, because trust rooted in transparency separates the news from other providers of information which, third, is increasingly critical for news organisations to sustain their business.

Transparency in news has multiple definitions and for the purpose of this paper, we will mostly focus on disclosure transparency – "various techniques illustrating to the public how and why the news is being made, by, for instance, explaining the news selection process and so opening up news production for insights" (Karlsson, 2020). In practice, disclosure transparency includes such rituals as the use of bylines, citing sources in texts and graphics, explaining the methodology of reporting and the process behind it, or issuing corrections of errors. The internet has made some forms of transparency more interactive and real-time: for instance, offering original datasets for download or allowing for more extensive bylines.

However, further advances of technology in the forms of social media and AI have resulted in both growing demand for transparency and new practices – but also challenges – to transparency norms and disclosure practices (Cools and Koliska, 2024). On the one hand, the news is part of a large technological system (Simon, 2025) in which decisions taken by other actors can affect journalistic work and credibility (e.g. the way news is displayed in search engines). On the other hand, AI systems can reduce the autonomy of newsworkers because they make decisions or produce output (semi-)independently. Both developments have complicated questions of when, how, and to what extent the news has to be – and can be – transparent about how "the sausage gets made".

Audience responses to AI use and AI transparency in news

Audiences seem well aware of some of these difficulties and hold nuanced and often sceptical views on the use of the technology in the news. Recent empirical work has found that audiences across countries have limited trust in journalists and newsrooms to use these systems responsibly (Mitova et al., 2025, Simon et al., 2025), expect news produced mostly by AI to be less trustworthy and less transparent (Fletcher & Nielsen, 2024, Newman et al., 2025), are cautious about AI-use for personalisation (Newman, 2025), and want news organisations to be transparent about their AI use (Fletcher & Nielsen, 2024; Vogler et al., 2023; CNTI Briefing 1, 2025).

In response to these concerns and newsworkers' worries, many media outlets and industry organisations have sought to develop responsible AI practices, which include a focus on transparency. For example, many publishers produced public statements and internal guidelines

about how they plan to use AI responsibly. Transparency emerges as a recurring theme in many of these guidelines (de-Lima-Santos et al., 2024; Becker et al., 2025). The main goal often seems to be to address potential errors that AI systems might make upfront, with some organisations explicitly communicating the reliability of these systems. Multiple approaches currently exist to further transparency around AI in news: (1) technical openness (such as explainable AI and algorithmic audits); (2) process transparency (information about the design, development and deployment decisions and practices behind AI solutions); (3) outcome-based transparency and explainability (clarifying to users exposed to a service that utilises AI which factors influence its decision making and output); (4) public transparency (clear communication about when and how AI is used in news production and distribution) (Diakopoulos, 2019; Government Digital Service, 2025).

Yet, various challenges in implementing these approaches have been found. Practitioners often lack the data and/or algorithmic literacy to understand how AI systems work, and some systems are hard or costly to interpret even for their creators, making it difficult to explain their workings. Transparency can also be limited by proprietary and 'black box' commercial tools. Furthermore, simply providing explanations is insufficient if users cannot understand or interpret them correctly, especially if the explanations are too technical or potentially misleading (Dierickx et al., 2024; Government Digital Service, 2025). There is also no shared consensus on what transparency should look like in practice, and at what stages. Research has also found that transparency about AI can backfire, reducing trust in news (Toff & Simon, 2024).

Research Questions

While existing research covers high-level principles of AI transparency in news, audience attitudes towards the same and some of the challenges tied to it, a gap remains in understanding specific newsroom practices related to AI transparency, especially how organisations are approaching and implementing the same and how they balance ethical concerns around AI and a desire for transparency with other existing organisational constraints and processes. To better understand this dynamic, this paper asks:

> **RQ1**: *How does a leading news organisation integrate AI transparency into their practices?*
> **RQ2:** *What factors shape the disclosure and communication of AI use to audiences and newsworkers?*
> **RQ3:** *What are challenges around AI transparency?*

Methodology

In order to address these questions, we employed a case study approach focused on the Financial Times (FT) – a leading global news organisation that has actively been integrating AI into its newsroom for several years. To study this topic in depth, this research relies on two primary sources of data: (1) semi-structured interviews with newsroom leaders from different departments; (2) internal materials and documents relating to the FT's AI use.

Selection of News Organisation

The Financial Times (FT) is a globally known business news publisher founded in 1888 and in recent years developed a strong reputation for digital innovation. This includes the adoption of AI. For example, the FT has (1) deployed a machine learning comment moderation systems to assist moderators and boost the proportion of constructive reader contributions (Kunova, 2024); (2) launched "Ask FT," a generative AI chatbot that lets subscribers interrogate almost two decades of archival reporting (Roth & Kennemer, 2024); (3) introduced AI-driven related article recommendations through its FT Professional service (Financial Times Professional, 2024); and is currently trialling (4) real-time research and summarisation assistants (Financial Times Live, 2025; Reuters Institute for the Study of Journalism, 2025), among other things. The FT also provides expertise to other publisher via FT Strategies, its consulting arm. The FT was one of the first news organisations to sign a content licensing partnership with OpenAI (Milmo, 2024). Coupled with the unparalleled access and openness extended to the lead researcher – enabling direct observation of product, editorial and data teams – these factors make the FT a rich and timely case through which to interrogate in-depth how an established news organisations adopts, governs and operationalises transparency around AI.

Data Collection

Semi-structured Interviews

Following a strategic, purposive sampling approach, we first identified professionals within the FT directly involved in AI adoption and transparency practices in the organisation using the lead researcher's existing knowledge of the FT's hierarchy and desk-based research of AI activities at the FT. After receiving internal approval, the lead researcher approached and interviewed seven newsroom leaders working across editorial, product, and data science to capture diverse perspectives on the topic with two rounds of interviews between end of May 2025 and mid-July 2025, with the initial group of interviewees and desk-research leading to the participants in the second round. In total, we conducted 13 interviews with 12 participants (one participant was interviewed twice). A list of all interviewees can be found in Figure 1.

We used the same semi-structured interview instrument in both waves to address the research questions in a standardised format, while providing flexibility to discuss aspects which arose during the interviews. We formulated interview questions based on the existing literature and the research questions to address: (1) the current AI use in news production and product development at the FT; (2) decision-making processes around internal and external communication, as well as transparency and disclosure of AI use; (3) perceived audience expectations and reactions to AI in journalism. All interviews lasted on average 30 minutes and were recorded and transcribed with participant consent. We used the FT's internal AI transcription service for data security reasons and transcription quality. We took detailed notes and wrote brief reflective memos after each interview which formed part of the final data set for analysis.

ID	Gender	Seniority level	Function	Role in the context of AI	Date of the interview
I1	Male	Senior Leadership	Public Policy	Industry collaboration on issues related to AI	01/07/2025
I2	Male	Manager	Editorial	Collaboration with product and tech teams on AI capabilities for reporters; member of the Editorial AI Use Case Panel	08/07/2025
I3	Female	Senior Individual Contributor	Product	Leading a cross-functional product and tech team focused on AI applications in the newsroom	01/07/2025
I4	Female	Manager	Internal Communications	Leading employee AI adoption and communication, including messaging around internal AI experimentation ahead of subscriber launches	26/06/2025
I5	Male	Senior Leadership	Editorial	Collaboration with product and tech teams on internal and reader-facing editorial products, including AI-powered; member of the Editorial AI Use Case Panel	02/07/2025
I6	Male	Senior Individual Contributor	Product	Leading a cross-functional product and tech team focused on content personalisation and search experience	02/07/2025
I7	Female	Senior Individual Contributor	Data Science	Leading data science work for GenAI solutions	16/06/2025 & 08/07/2025
I8	Female	Senior Leadership	Product	Leading all product development across the FT, with the focus on reader experience	23/05/2025
I9	Male	Senior Leadership	Editorial	Leading digital platforms in the newsroom, with the focus on AI initiatives; member of the Editorial AI Use Case Panel	25/06/2025
I10	Female	Senior Leadership	Editorial	Leading collaboration between product and tech teams and the newsroom	25/06/2025
I11	Male	Individual Contributor	User Research	Qualitative user research, currently on the team focused on AI applications in the newsroom	25/06/2025
I12	Male	Senior Individual Contributor	Data Analytics	Quantitative data analysis and experimentation, currently supports the team focused on AI applications in the newsroom	26/06/2025

Internal Materials

In addition to the interviews, this study examined internal FT materials, including guidelines, policies, internal reports, and meeting minutes related to the topic. These materials provided contextual background on transparency policies and the extent to which AI use is documented or standardised within the newsroom.

Data Analysis

Interview transcripts and internal materials were analysed using a mixture of deductive and inductive thematic coding, organising the data into key themes related to AI transparency and disclosure practices. We devised a coding scheme from the research question and the existing literature as well as from an initial read of the transcripts and documents. This coding scheme was then discussed and refined and then applied to all the data. As a last step, we thematically categorised codes – grouping similar codes into broader themes – and finally jointly interpreted and synthesised the data.

Ethical Considerations and Positionality

The research received ethical approval by the University of Oxford's Central University Research Ethics Committee (CUREC reference: SSH/DPIR_25_1650874) and was internally approved by the Financial Times' management. Participants were provided with an information sheet and a consent form. We guaranteed anonymisation by default to protect participants and encourage open discussion and have anonymised participants throughout. The Financial Times received the right to review only the final draft to ensure the protection of proprietary or sensitive information but otherwise had no influence on the framing, conduct, or final version of this study.

We also note our own positionality. As the Product Director responsible for AI at the Financial Times at the time of writing, the lead researcher's position informs this research in two ways. First, their insider perspective provided us with a deeper knowledge of the FT's structures and procedures and allowed us to develop a nuanced understanding of the transparency challenges around AI within the FT. More importantly, it also allowed us to gain unrivalled access to an elite setting (Harvey, 2011) and various key personnel at the FT and was key in building rapport with participants. While this familiarity with the FT was beneficial, it also introduced the potential for bias. To mitigate this as much as possible, we avoided leading questions during the interviews and discussed findings and assumptions with the external researcher who is not embedded in the day-to-day operations of the newspaper. We also drew on wider literature to triangulate and contextualise the findings.

Findings

Motivations for AI transparency at the *Financial Times*

The communication of AI use and transparency about the same is driven strongly by a desire to maintain trust and the FT's reputation for accuracy – something that was highlighted by all our interviews. Transparency about AI was seen as a key safeguard and form of reassurance to maintain both reader trust and internal trust. However, in our interviews it emerged that this presented slightly differently depending on the imagined – internal or external – audience.

Internal open communication and transparency around AI use was framed across interviews as essential to reduce frictions, ultimately enabling smoother workflows and better integration of AI because people can learn from each other about successes and failures. As one manager (I4) put it: *"We've talked about decisions where we tried something with AI and it didn't work, and we've been very transparent about that as well as the times when [it did work]. I think that really helped to move the dial."* Being transparent in the internal communication and disclosure of AI use was also described as needed to reduce the risks stemming from the use of the technology and catching any errors – either in the way systems function or in the output they create – as early as possible, before the same could reach readers or lead to worse decisions. Finally, transparency about AI use and open communication was described by many managers we interviewed as central to reducing fears around the technology, which were seen as ever present, at times for good reasons: "*When you job is about creating words and people are talking about using a tool that creates words of varying quality, you are understandably very cynical and concerned about it"* (I4).

Meanwhile, *external* communication and transparency around AI focuses on maintaining the outlets' reputation as a reliable, high-quality news brand that can be trusted as such. In the words of one senior manager (I10): *"It all ties back into trust [...] because it's core to what we're doing"* with another (I4) remarking that "*as a news organisation that holds other people to account we have to make sure that we walk that walk.*" Here, too, transparency about AI use also entailed an element of risk reduction, namely by being able to at least somewhat control or direct reader reactions. *"It's trying to manage the expectations of readers"*, as one participant (I7) put it. Maintaining trust in an audience-facing context, however, was not merely described as a normative principle in line with journalisms' professional rituals and ethics, but also as a competitive asset and necessity to stay in business. Several interviewees, as well as the FT's documentation and public branding, consistently underline that audiences' trust in the Financial Times delivering a high-quality "news product" was what made people willing to pay – and thus makes the outlet commercially successful. Or in the words of interviewee I10: *"Mistrust is the top threat. You lose that [trust], forget the rest."*

AI Transparency as a combination of proceduralism, leadership endorsement, and experimentation

A core theme from the data is that the FT's approach to transparency around AI is enabled and enforced top-down and structurally anchored in a set of external and internal guidance documents

as well as procedures. The FT's commitment to AI transparency is endorsed by its senior leadership – something that was confirmed by both our interviews and the documents we analysed. Both Roula Khalaf, the current Editor in Chief, and John Ridding, former CEO, stated their commitment to AI transparency when addressing readers and internal staff in their communication in the past:

> "At a time when misinformation can be generated and spread rapidly and trust in the media in general has declined, we at the FT have a greater responsibility to be transparent, to report the facts and to pursue the truth. [...] (Ridding, 2023)

> We will be transparent, within the FT and with our readers. All newsroom experimentation will be recorded in an internal register, including, to the extent possible, the use of third-party providers who may be using the tool." (Khalaf, 2023)

One of the internally published AI Principles also reads:

> "We will not allow AI to compromise the integrity of our journalism which will be reported, written and created by our journalists and editors. We will experiment with AI to create secondary article content (e.g. summaries, alt-text, captions, translations) as long as [...] there is appropriate disclosure".

This is already an example of committing to transparency: openly stating FT's position on AI both to readers and employees.

The main process behind enforcing this approach is the review of new potential AI applications by AI Use Case Panels whose goal is to make a decision about approving a specific use case and advise on its implementation – including but not limited to the appropriate level of transparency and disclosure. The internal FT's AI Policy states:

> "AI should be used in an ethical and responsible way; this means considering the potential impact on individuals, society and the environment. The AI Panel will consider the following points when assessing submissions, and we encourage all users of AI to do the same: [...] Transparency – not concealing a use of AI or its workings; this usually involves disclosure".

While the process is to review individual cases and evolve guidance over time, there are some red lines that the FT stays within – for instance, around AI-generated images. As one interviewee (I9) described it:

> "We haven't used generative AI to produce imagery. We've only done that a couple of times to illustrate stories about AI. For instance, to give an example when covering an argument about copyright theft over AI. The images in such

> articles are clearly flagged. We're not producing AI images and trying to slip them into stories."

For cases when the line is "orange, rather than red" (I8) to find a balance between opportunities of AI and risk associated with it, the FT prefers an experimental approach – carefully engaging with the new technology, observing audience and newsroom reactions in a controlled way to gather data and inform decisions. One interviewee (I1) highlighted the value of FT staff *"having the freedom to play around with things like ChatGPT[1]"* as a way to better understand the new technology and what it is capable of – which in turn informs transparency decisions. An editorial leader (I5) asserted that *"there is no point in asking questions in theory and therefore it is really critical that we can experiment with small groups, ask for feedback, see the data"*. Representatives of the data function in our interviews were unanimously underlying the importance of relying on experimentation data when making decisions about rolling out AI capabilities and transparency decisions that go along with it:

> "The opportunity to test new features with FT employees first is a great advantage: our diverse staff are skilled in multiple languages, so they can give us good feedback on GenAI output quality, especially for features like translations." (I7)

With the evolution of AI capabilities and more in-house experimentation, our interviewees reported that the confidence about LLM's accuracy and reader behaviour in reaction to new features has grown.

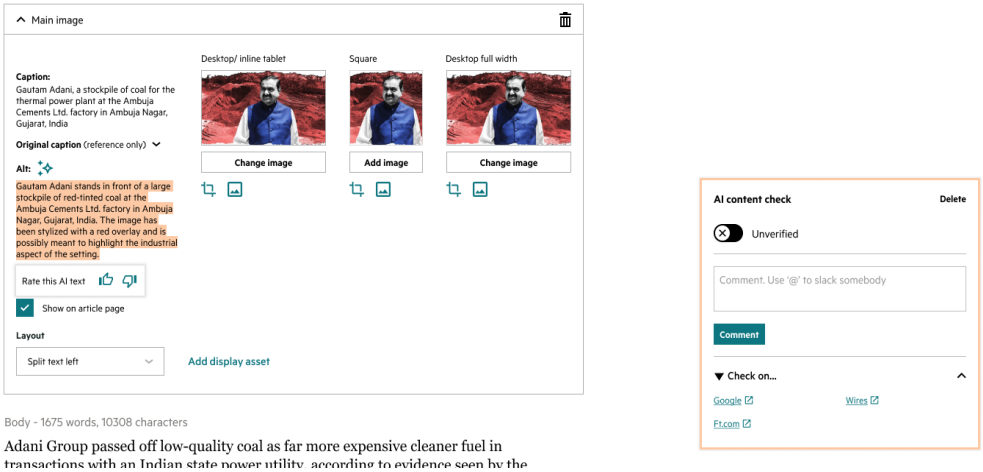
Internal AI Transparency: Communication, Design and Training

Internal transparency at the FT is aimed at reducing friction, addressing anxieties from staff around AI, and preventing unauthorised AI use in workflows. The internal AI Policy places accountability on staff to apply AI only where appropriate and to review its output "as if you created it yourself," requiring clear signposting in day-to-day tools. I9 explains: *"Reporters know they're still responsible: if an AI-generated summary has an error, their name is on the line just as before."*

Transparency to internal users is being achieved via design and training. *"Clear internal training on AI tools allows journalists to understand when AI is being used and how much they should trust it"* (I7). In some critical internal tools like the main content management system, the FT follows disclosure practices like those to the external readers – adding reminders that AI-generated content may contain hallucinations and signposting AI functionality with the "sparkle emoji". Additionally, there are technical efforts at the FT to highlight parts of the AI output that are more likely to contain errors and require closer human attention.

[1] ChatGPT – an LLM-powered conversational agent developed by OpenAI.

Figure 4: Example of an AI-assisted feature in the internal content management system



Communication around AI aims to be targeted and structured. As one interviewee explained, *"the Intranet serves as a single source of truth with screenshots and other multimedia formats to aid clarity,"* (I4). She also told us about the AI Transformation group which acts as a forum *"to avoid siloing of information and to make sure that when someone has something that they've learned that's shiny and new or a little bit worrying, they can bring it up in that forum and it then gets the right kind of people."* Additionally, specific internal user groups are reached in familiar forums and by leadership representatives, with the aim of having open discussions. Openness is reinforced through regular sharing sessions across editorial, product, and tech teams. *"Right up to very senior leadership there is a willingness to put hands up and go, 'We don't know',"* I10 observed, emphasising a cautious, methodical approach also highlighted by other interviewees with "a lot of data and cheques... and kill switches the whole way along." This willingness to admit uncertainty and proceed carefully was seen by participants as one of the most transparent aspects of the FT's AI projects.

Nine factors inform the "What? When? How?" of reader-facing disclosure of AI

When it comes to applying AI transparency in reader-facing context, the FT employs a collaborative approach that considers multiple inputs and points of view. We identified nine factors that contribute to the level of transparency and shape of disclosure of AI use to readers.

First, **legal requirements** are an important shaping factor. Some AI providers contractually require disclosure when their model is being used in particular circumstances. For instance, the AskFT[2] feature includes the following disclaimer: *"This tool generates automated answers summarised from FT articles. By using it, you are interacting with an AI system (provided by Anthropic), not a human"*. AI providers' motivations for this were not clear to the participants of this study, but interviewees speculated that a combination of gaining visibility and protecting themselves legally

[2] AskFT – an interface on FT.com allowing readers to ask a question in natural language and receive an answer generated by an LLM, based on the archive of content published by the FT.

for cases of hallucinations may be some of the reasons. As I7 explained: *"I may even imagine different motives, like alerting of what is going on and how people are using their model"*. While some providers only state high-level requirements, others are very prescriptive. I7 shared an example: *"For Google Translate, a very specific logo image was necessary to include at the start of the translated text, as well as a paragraph of capitalised text at the end of it."*

Second, we found that the FT tends to do **industry benchmarking** and align its approach with other leading publishers, following industry norms for AI labelling and disclosure. Participants mentioned brands such as the New York Times, the Washington Post, or Aftonbladet (among others) which have live AI features to learn from as well as the publicly released AI Transparency guidelines of organisations like the BBC. In the words of I3: *"We looked at a ton of external examples, and we are in line with the way that most other publishers have approached these same sorts of considerations*". Best practice insights which according to our interviews shape their decisions are also gathered at industry events frequently attended by staff involved in AI development.

Third, the **nature of the tasks or procedures affected** affects decision-making. Some behind-the-scenes AI-driven processes in the newsroom do not affect content directly and because of it normally do not require disclosure. *"We are not disclosing AI use by the SEO[3] team because its output is not immediately public-facing. It is just assisting in workflow processes"* (I10).

A fourth factor is the **level of human oversight required**, with AI-generated content edited by journalists treated differently from fully automated outputs: *"AskFT is completely unassisted and has no human in the loop, so the disclaimer is stronger than the Bullet Point Summaries[4]"* (I3). This approach is reflected in the FT's internal Responsible AI Guidelines:

> "We aim to have an appropriate level of transparency to ensure trust and understanding, particularly in any external facing use of AI. The more automatic or impactful a process, or the larger the output proportion generated by AI, the more appropriate it is to be transparent and the more prominent disclosure should be. Ensuring that we let people know when their data is being handled by AI, and when what looks like something we are doing directly is actually something done by AI ensures trust and clarity in the relationship. When in doubt, disclosure is the default option".

The higher the level of autonomy of a system, the more likely that strong transparency and disclosure would be required, according to our interviewees. In addition, we found that participants felt that the higher the contribution of AI to a final product is, the higher the need to disclose its use became.

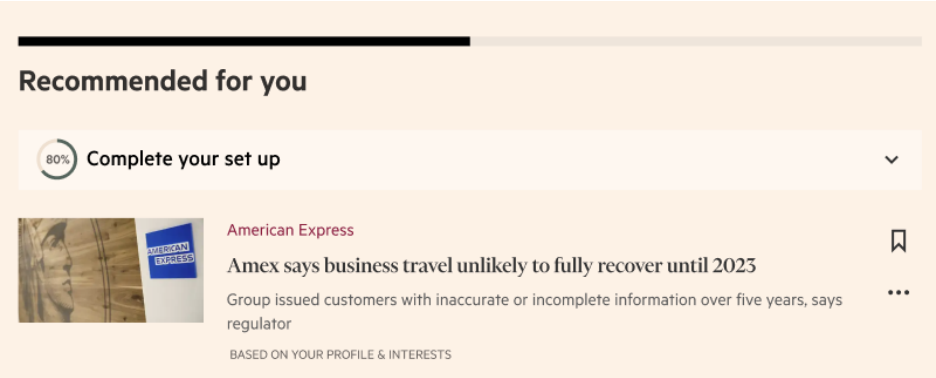
[3] SEO (Search Engine Optimisation) – the practice of improving a website's visibility in search engine results pages to increase organic (non-paid) traffic.

[4] Bullet point summaries – an element in some FT news articles with four bullet points that summarise the content, generated by an LLM and not reviewed by an FT journalist before publishing.

A fifth factor is the **newness of AI systems** and therefore lower levels of practical experience and greater uncertainty around their performance. As I8 argued, the question *"Are we using an existing tool and existing frameworks or are we trying to use something new?"* plays an important role in decisions about transparency. At the time of writing, this factor was particularly salient for many interviewees given the staged roll-out of a new feature in which the FT decided to "take the human out of the loop" for a reader-facing application with direct impact on news content: *"There is quite a departure in saying, for the first time, we actually take the human out of the loop—and we'll make that known to readers"* (I10).

Sixth, **audience expectations, audience understanding, and (real or anticipated) audience reactions** are key factors for audience-facing disclosures and transparency. To gauge audience views, the FT employs usability testing[5] as a method to understand reader perceptions of new functionalities and conducts various forms of audience research both pre- and post-deployment of new features. Supported by such user feedback, the FT adopted the "sparkle emoji" (✨) as a visual signpost of AI-generated content: it is understood by readers and is also becoming an industry-wide standard. Audience expectations also shape the choice of colour: "*We had debates about the colour of the background... to show that it's not the journalism; it's something on top of it."* (I8) While there is no universal FT guidance on whether such use of AI needs to be disclosed, reader reactions play a shaping role. *"Users want to know why they are getting a recommendation – removing that 'black box' feeling is important"*, as I6 explained. This increases in magnitude when the suggestion is either "bad" or "too good".

Figure 3: Example of disclosure of factors behind AI-powered recommendations

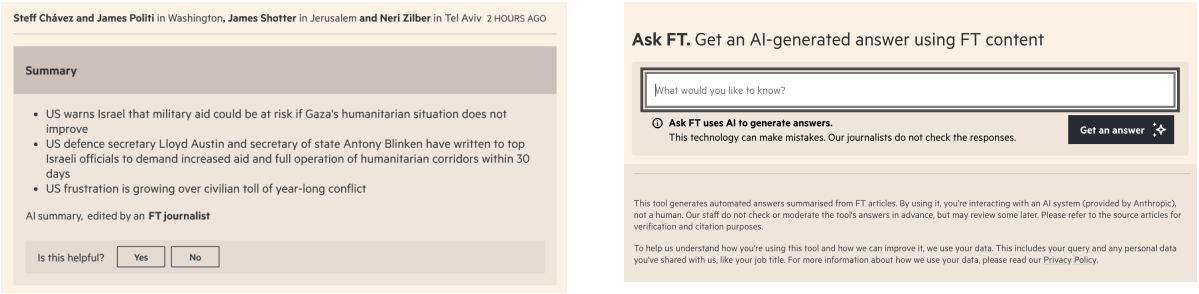


The seventh factor, strongly connected to the fourth and fifth, is the **perceived risk level of an AI application, both in terms of its performance as well as audience attitudes towards it**. Regarding the latter, internal and external research the FT leans on shows that readers see some uses of AI as more misleading, especially where they come closer to human likeness, such as in the case of voice cloning with such uses, which the FT hasn't explored, would warrant disclosure. Similarly, participants were aware that AI disclosure can backfire and make audiences question content, so avoiding over-labelling was seen as important: *"There is a tension between the ethical*

[5] Usability testing – a research method used to evaluate how easy and intuitive a product, system, or interface is for users. It involves real users interacting with a product while observers or researchers analyse their behavior, feedback, and difficulties.

obligation to disclose AI use and the risk of undermining trust in the FT brand" (I2). Regarding the risk level, interviewees described that they are particularly concerned about insufficient disclosure of AI use where there is a higher risk of failure, for example, because it provides inaccurate content. Apart from disclosure, placing AI-generated content in new elements which are more easily distinguishable from writing by journalists was described as a key mitigation.

Figure 2: Examples of AI use in content creation. On the left: Bullet point summaries (human review; location: article page). On the right: AskFT (no human review; location: bespoke page).



Eigth, **commercial constraints** play a role too. For instance, when AI's role in an article production process was to enable a complex investigation through data analysis, disclosing how exactly AI was used may be extremely costly because of the time it takes to create the necessary documentation. I2, for example, explained that describing AI methods is not a straightforward task: *"You could do all of the nerdy stuff on GitHub[6]"*. Separately, such disclosures could put some of the FT's competitive edge at risk – for instance, when giving up unique data sets used to uncover a story.

Finally, nine, **design constraints** which often intersect with commercial considerations, play a role. Disclaimers can take up space on the website that could be devoted to promotion of other materials or create a less clean presentation that depreciates the experience of a high-quality product as audience views develop: *"To clutter the page in our features with tiny text explaining what we've done and what model we're using becomes less and less helpful because it will become intuitive for the user [...]."* (I10) Meanwhile, lengthy AI disclosures could be at odds with limitations on articles lengths, as on interviewee argued.

Four Challenges for Internal and External AI Transparency

Finally, we identify four challenges around AI transparency. First, the **consistency and design of AI disclosure** is seen as an area for the FT – but also the news industry and other industries such as streaming – to invest in as AI use expands. *"I think there are discussions within industry about what proper attribution and labelling looks like"* (I1), said one interviewee but several participants argued that while a consensus around this topic was building within the news industry, there were still no generally agreed upon approaches. More importantly, this was seen as challenging internally, too. In the words of another participants, *"We are missing a unified FT-wide agreement*

[6] GitHub – a cloud-based platform for version control and collaboration that allows developers to manage and share code.

on design components and transparency guidelines. Readers will trust us more if our AI disclosures are consistent across the site" (I3). One issue the FT has identified is to create clear disclosures on smaller screens: *"Users find [what is and is not AI-generated] a lot more obvious on desktop than they do on mobile because you have more real estate on a desktop experience to actually call out specific elements"* (I3).

Second, **breaking organisational silos and effectively communicating** both the internal and external use of AI in a way that it a) reaches busy and/or AI-sceptical newsworkers, b) is accessible and easy to understand and c) "sticks" with them (in the words of one interviewee) emerged as a recurring challenge in being transparent about the use of AI, especially against the backdrop of time and resource constraints. As one interviewee (I4) put it: *"We have the responsible and ethical use of AI framework which does give you quite specific guidance [...]. I don't think you're going to find it's more than a couple hundred people out of three and a half thousand [who have read it]."*

Third, an **evolution of standards with AI advances** was seen as a future challenge. AI tools are improving at the same time as the public is getting more used to their benefits and limitations, resulting in a need to continually reassess transparency expectations and norms, according to our interviewees. As I5 put it: *"AI guidance today might not apply next year. At what point does AI become just another editorial tool and not something that needs to be labelled?"* Another interviewee (I10) drew a comparison with entertainment platforms:

> "It may be a silly example, but I'm just thinking of like the Spotify's and the Netflix's of the world. We know that they've got these models running. We absolutely know, there's no way you cannot notice it, but you're not getting a pop up warning you. And so, I think we are just much, much, much earlier on the journey in terms of how news is handling that.

Fourth, and relatedly, **adapting not just to AI advances but to changes in human behaviour** was also predicted to be a necessity. Even though newsworkers – according to our participants – were currently very cautious and diligent in reviewing AI outputs, one participant argued that we should expect a greater *"tendency for humans to rely on these tools over time, especially once they gain confidence in them"* (I2), with internal and external AI transparency tailored accordingly.

Fifth, and finally, **audience expectations and audience understanding of AI transparency can be complex.** For example, the FT realized that the presence of AI disclosures may cause some confusion for readers, as one participant explained: *"A few participants assumed the presence of the AI tools indicated that the article itself had been generated by AI"*. However, this was seen as context-dependent, with I2 expressing the belief that for the use of AI in investigations *"the mere fact that disclosure of sources and methods behind an investigation is there increases the reader's trust in the rigor of it"*.

Discussion: Tricky Trade-Offs Amongst Competing Logics

Like other media organisations, the FT views AI transparency as a critical ethical issue and business challenge. The findings of this study support previous research and add a new level of granularity to previously covered approaches to AI transparency in a news media organisation.

In answer to RQ1, we find that the FT integrates AI transparency through a hybrid of policy, process, and practice, underwritten by explicit senior-leadership endorsement and a risk-sensitive, experimental posture. Transparency is framed as instrumental to safeguarding trust – both internally (to reduce frictions, surface errors early, and lower anxiety among newsworkers) and externally (to reassure readers and protect the brand). Operationally, the FT anchors transparency in both formal artefacts (e.g. AI principles, cross-functional AI Use Case Panels) and in day-to-day design choices and training. Internally, full disclosure is the stated aim and any AI use is signposted in tools, accompanied by reminders about limitations and personal accountability, and supported with targeted communication and skills development. Externally, the strength and prominence of disclosure are calibrated to system autonomy and editorial oversight: fully automated, no-human-in-the-loop features (e.g., AskFT) receive stronger and more prominent labelling than AI-assisted, journalist-edited outputs (e.g., bullet-point summaries). Across contexts, the FT's "experiment first, measure, and iterate" approach supplies the evidence needed to refine both guidance and practice over time.

Addressing RQ2 and RQ3, we identify nine factors that systematically shape what, when, and how the FT discloses AI use to readers: (1) legal and provider requirements; (2) industry benchmarking; (3) the degree of human oversight and proportion of AI contribution; (4) the nature of the task affected; (5) the novelty and uncertainty of the system; (6) audience expectations and audience research findings; (7) the perceived risk of harm or error; (8) commercial sensitivities (costs of method disclosure, competitive exposure); and (9) design constraints (space, page cleanliness, device form factor). These drivers sit alongside five cross-cutting challenges: establishing consistent, site-wide labelling (especially on mobile); overcoming silos and communication fatigue internally; keeping pace with shifting norms as models – and user familiarity – evolve; guarding against growing human over-reliance on tools even when labelled; and mitigating audience misinterpretations or "transparency backfire."

This paper does not try and present a new theory of AI or transparency in the news. Instead, through thick description and close observation we have sought to identify how AI transparency is enacted in news practice, what factors shape it, and what challenges arise. We argue that these findings complement and provide new evidence for existing theories and arguments of how digital innovation plays out in the news – arguments which we will outline below.

AI Transparency is intertwined with isomorphic pressures and institutional logics

The first argument relates to institutional theory which holds that (news) organisations are influenced in what they do by the environment they operate in and that they model themselves at least partially on other, similar organisations to deal with uncertainty and reduce risks

(DiMaggio and Powell, 1983). This has been observed in previous research for their approach to AI (Becker et al., 2025; Hilker et. al., 2025, Simon, 2023), and we argue that *the way transparency around AI is approached and implemented at the FT is shaped by such isomorphic pressures*, too.

This becomes clear with a look at the factors shaping AI transparency: Regulations and the laws that demand transparency around AI can be seen as *coercive pressures* – "[...] pressures exerted on organisations by other organisations on which they depend and by cultural expectations in the society in which organisations operate" (DiMaggio and Powell 1983, 150). In a similar vein, audience expectations that the news (and the Financial Times as a news organisation) should be transparent about its AI use should be interpreted as a form of coercive pressure. *Normative pressures* – "the collective struggle of members of a profession to define the conditions and methods of their work" (DiMaggio & Powell, 1983, 152) – are evident here too, with our findings showing that the FT adopts transparency measures because it is something done by others, too, and because they have become an industry norm. Finally, the exact approach to AI transparency taken by other leading publishers also shapes the newspaper's approach in an area where single, correct answers do not exist, uncertainty reigns because audience expectations are complex, and the issue at hand is sufficiently complex. The uncertainty of the best approach leads to *mimetic pressures* and the adaption or copying of approaches seen elsewhere.

Our second argument is that *AI transparency cannot be seen as separate from the institutional logics that a news organisation like the Financial Times – implicitly and explicitly – adheres to.* We understand institutional logics here as "systems of cultural elements (values, beliefs, and normative expectations) by which people, groups, and organisations make sense of and evaluate their everyday activities, and organize those activities in time and space" (Haveman & Gualtieri, 2017). Logics are also sometimes described as prescriptions that shape what is considered legitimate or desirable in a field (Hilker et al., 2025, 5). What this boils down to is that, in practice, AI transparency at the FT reflects multiple intersecting logics. Internal transparency decisions are primarily aimed at reducing friction, managing employee anxieties, and preventing unauthorised or improper use of AI systems, thereby supporting the adoption of AI and mitigating operational risks. These practices exemplify a managerial logic (see also Thornton et al., 2012), which emphasises efficiency, the optimization of workflows, and the rational organisation of news work (Simon, 2025). This managerial logic is tied into a commercial logic, in which news production is understood as a competitive product in a crowded media environment beset by competition for limited attention (and by extension the purchasing power) of information consumers (Jungherr & Schroeder, 2021).

Externally, AI transparency decisions serve to manage the organisation's reputation and maintain audience trust. Here, the professional logic of journalism is evident, where transparency signals to readers that reporting – and the organisation producing it – are accurate, accountable, and trustworthy (Deuze, 2005). Simultaneously, a commercial logic operates externally as well, as trust fostered through transparency about AI differentiates (or at least so is the hope) the FT from other information providers, thereby (ideally) enhancing its perceived market value in the eyes of the audience. In economic terms, this form of transparency is ultimately a form of expectation management, where signalling reliability and quality influences consumer perceptions and purchasing behaviour.

AI transparency is a lived practice shaped by a spectrum of competing logics

A final point is that all things considered, our case suggests *AI transparency in news organisations is best understood as a situated, continuously updated and lived governance practice* – one that balances ethical commitments with organisational and commercial realities rather than a one-off, one-size-fits-all rule or binary. This observation is also one made tacitly by our participants themselves who hint at these tensions and the often tricky trade-offs involved, even if they do not expressly articulate the same. It also falls in line with recent audience research (Fletcher & Nielsen, 2024; Morosoli et al., 2025) which finds that audiences see – and demand – AI transparency not as a binary black and white issue but changing from context to context. As such, *we argue that AI transparency might be best understood as a practice sitting on a spectrum of sometimes competing logics* and recognised as such by both newsworkers and audiences. Competing, because optimising for one factor (e.g. full transparency as demanded by regulations) might come at the expense of audience trust or revenue, where transparency makes audiences distrust what they consume (Toff & Simon, 2024; Zoizner et al., 2025). For now, the idea of transparency as a spectrum-shaped practice is only a hypothesis. Further research will need to investigate not only the generalisability of the shaping factors and challenges we describe beyond a single organisational context, but also address if there are other factors or challenges not described here. It should also investigate if there is something like an optimal balance between these different factors in terms of both serving the organisations' goals and audiences' needs.

Conclusion

This case study shows that AI transparency at the Financial Times (FT) is neither a single rule nor a binary label but a situated governance practice that is continually adjusted to technology, audience expectations, and organisational constraints. The FT operationalises transparency through a hybrid of policy, process, and practice. Internally, transparency is integrated into tools and workflows (e.g., signposting functionality, training, and explicit accountability), primarily to reduce friction, surface errors early, and build confidence. Externally, the strength and placement of disclosure scale with AI system's autonomy and editorial oversight, with no-human-in-the-loop features (e.g., AskFT) labelled more prominently than AI-assisted, journalist-edited outputs (e.g., bullet-point summaries).

We identify nine factors that systematically shape the "what, when, and how" of AI transparency and which can lead to tricky trade-offs for the *Financial Times*: (1) legal/provider requirements; (2) industry benchmarking; (3) degree of human oversight and proportion of AI contribution; (4) nature of the task; (5) novelty/uncertainty of the system; (6) audience expectations and research; (7) perceived risk of harm or error; (8) commercial sensitivities; and (9) design constraints. These levers coexist with persistent challenges: ensuring consistent labelling; communicating effectively across organisational silos; updating norms as models and newsroom practices evolve; guarding against creeping human over-reliance; and mitigating "transparency backfire" and audience misinterpretation.

Theoretically, our findings illuminate how isomorphic pressures and intersecting institutional logics jointly structure transparency choices at the *Financial Times*. Practically, we suggest several "no-regret" principles that other organisations could adopt: To scale disclosure with

autonomy and impact; to pair AI labels with concise, comprehensible explanations; to measure user interpretation continuously; to keep internal transparency and accountability high even when external disclosure is light; and to favour consistent, reusable design components across products and platforms to avoid over- or under-labelling. As a single case study, this research does not aim to generalise findings to all news organisations. Instead, it provides a detailed exploration of AI transparency practices within a leading news organisation. Future studies may expand on this work by examining AI transparency across multiple media outlets or conducting comparative analyses. Future research should also explore how audience perceptions of AI disclosure shift over time and how newsrooms can refine disclosure strategies to enhance audience trust.

Ultimately, transparency about AI in news is at once simple and complex – simple in its ethical rationale, but complex in its execution amid competing goals and constraints. As a final aside, we should also point out that its success is not guaranteed. The boundary between journalism and adjacent information experiences – which organisations like the FT try to reinforce, including through transparency about their AI use – may be less salient to users than to practitioners and scholars. Likewise, trust in news flows not just from how transparent news organisations are but from a plethora of other factors, (Fawzi et al., 2021; Ross Arguedas et al., 2023) and audience acceptance of AI in news is shaped as much by everyday technology use and broader cultural narratives as by newsroom choices. Many of these elements news organisations cannot really control and even the best forms of AI transparency will likely have to bow to these forces that determine how and why people come to trust the news.

Acknowledgments

The authors would like to thank the team at the Financial Times, particularly Matthew Garrahan, for supporting this study. We would further like to thank the participants from the FT who volunteered their precious time to help us in understanding this issue.

Disclosure Statement

The authors have no conflicts of interest to disclose.

Funding

This research received no internal or external funding.

References

Becker, K. B., Simon, F. M., & Crum, C. (2025). Policies in Parallel? A Comparative Study of Journalistic AI Policies in 52 Global News Organisations. *Digital Journalism*, *0*(0), 1–21. https://doi.org/10.1080/21670811.2024.2431519

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., Beygelzimer, A., ... Liang, P. (2021). On the opportunities and risks of foundation models. arXiv. https://doi.org/10.48550/arXiv.2108.07258

Cools, H., & Diakopoulos, N. (2024). Uses of Generative AI in the Newsroom: Mapping Journalists' Perceptions of Perils and Possibilities. *Journalism Practice*, 1–19. 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 Studies*, *25*(6), 662–680. https://doi.org/10.1080/1461670X.2024.2326636

CNTI Briefing #1: AI Literacy & Communication. (2025). https://cnti.org/article/aij-wg-briefing1/

Curry, A. L., & Stroud, N. J. (2019). The effects of journalistic transparency on credibility assessments and engagement intentions. *Journalism*, *22*(4), 901-918. https://doi.org/10.1177/1464884919850387

Diakopoulos, N. (2019). *Automating the News: How Algorithms Are Rewriting the Media*. Harvard University Press.

Dierickx, L., Opdahl, A.L., Khan, S.A., Lindén, C.-G. and Carolina, D. (2024). A data-centric approach for ethical and trustworthy AI in journalism. *Ethics and Information Technology*, 26(4). doi:https://doi.org/10.1007/s10676-024-09801-6.

DiMaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, *48*(2), 147–160. https://doi.org/10.2307/2095101

Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., Dohle, M., Engelke, K. M., Hanitzsch, T., Jackob, N., Jakobs, I., Klawier, T., Post, S., Reinemann,

C., Schweiger, W., & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media – a literature review and framework. *Annals of the International Communication Association*, *45*(2), 154–174. https://doi.org/10.1080/23808985.2021.1960181

Financial Times. (2024, March 25). *Financial Times launches first generative AI tool for subscribers* [Press release]. https://aboutus.ft.com/press_release/financial-times-launches-first-generative-ai-tool

Financial Times Live. (2025, November 5–6). *Future of AI summit 2025*. https://ai.live.ft.com/

Financial Times Professional. (2024). *FT Professional introduces AI-powered relevant article recommendations*. https://professional.ft.com/en-gb/blog/ft-professional-introduces-ai-powered-relevant-article-recommendations/

Fletcher, R., Andı, S., Badrinathan, S., Eddy, K. A., Kalogeropoulos, A., Mont'Alverne, C., Robertson, C. T., Ross Arguedas, A., Schulz, A., Toff, B., & Nielsen, R. K. (2025). The link between changing news use and trust: Longitudinal analysis of 46 countries. *Journal of Communication, 75*(1), 1–15. https://doi.org/10.1093/joc/jqae044

Fletcher, R., & Nielsen, R. K. (2024). *What does the public in six countries think of generative AI in news?* Reuters Institute for the Study of Journalism. https://doi.org/10.60625/RISJ-4ZB8-CG87

Government Digital Service (2025). *Artificial Intelligence Playbook for the UK Government (HTML)*. [online] GOV.UK. Available at: https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html [Accessed 22 Feb. 2025].

Harvey, W. S. (2011). Strategies for conducting elite interviews. Qualitative Research, 11(4), 431-441. https://doi.org/10.1177/1468794111404329

Haveman, H. A., & Gualtieri, G. (2017). *Institutional Logics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190224851.013.137

Hilker, C., Schaetz, N., & Lischka, J. A. (2025). The logics of innovation. Institutional logics and isomorphic pressures in public service media innovation labs in the age of artificial intelligence. Journalism, 0(0). https://doi.org/10.1177/14648849251349911

House of Lords Communications and Digital Committee: Large language models and generative AI (No. 1st Report of Session 2023–24; House of Lords Communications and Digital Committee). (2024). House of Lords. https://publications.parliament.uk/pa/ld5804/ldselect/ldcomm/54/5402.htm

Jungherr, A., & Schroeder, R. (2021). *Digital Transformations of the Public Arena*. Cambridge University Press. https://www.cambridge.org/core/elements/digital-transformations-of-the-public-arena/6E4169B5E1C87B0687190F688AB3866E

Karlsson, M. (2020). Dispersing the Opacity of Transparency in Journalism on the Appeal of Different Forms of Transparency to the General Public. *Journalism Studies*, *21*(13), 1795–1814. https://doi.org/10.1080/1461670X.2020.1790028

Kunova, M. (2024, June 10). *Keeping the conversation clean: How AI helps the Financial Times moderate comments*. Journalism.co.uk. https://www.journalism.co.uk/news/keeping-the-conversation-clean-how-ai-helps-the-financial-times-moderate-comments/s2/a1148237/

Milmo, D. (2024, April 29). *OpenAI to use FT journalism to train artificial intelligence systems*. The Guardian. https://www.theguardian.com/media/2024/apr/29/chatgpt-openai-ft-journalism-train-artificial-intelligence-systems

Mitova, E., Blassnig, S., Strikovic, E., Urman, A., Hannak, A., de Vreese, C., & Esser, F. (2025). Exploring Public Attitudes Toward Generative AI for News Across Four Countries. *Journal of Quantitative Description: Digital Media*, *5*. https://doi.org/10.51685/jqd.2025.012

Morosoli, S., Resendez, V., Naudts, L., Helberger, N., & de Vreese, C. (2025). "I Resist". A Study of Individual Attitudes Towards Generative AI in Journalism and Acts of Resistance, Risk Perceptions, Trust and Credibility. *Digital Journalism*, 1–20. https://doi.org/10.1080/21670811.2024.2435579

Newman, N., Fletcher, R., Robertson, C. T., Arguedas, A. R., & Nielsen, R. K. (2025). Reuters Institute Digital News Report 2025 (Digital News Report). Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025

Nielsen, R. K. (2017). The One Thing Journalism Just Might do for Democracy. *Journalism Studies*, *18*(10), 1251–1262. https://doi.org/10.1080/1461670X.2017.1338152

Official Journal of the European Union (2024). *Artificial Intelligence Act*. Available at: http://data.europa.eu/eli/reg/2024/1689/oj [Accessed 22 Feb. 2025].

OpenAI. (2023). GPT-4 technical report. arXiv. https://doi.org/10.48550/arXiv.2303.08774

Reuters Institute for the Study of Journalism. (2025). *AI and the Future of News 2025* [Video]. YouTube. https://www.youtube.com/watch?v=ao4JsYSyrD0

Ross Arguedas, A. A., Badrinathan, S., Mont'Alverne, C., Toff, B., Fletcher, R., & Nielsen, R. K. (2023). Shortcuts to trust: Relying on cues to judge online news from unfamiliar sources on digital platforms. Journalism, 25(6), 1207-1229. https://doi.org/10.1177/14648849231194485 (Original work published 2024)

Roth, E., & Kennemer, Q. (2024, March 23). *Financial Times tests an AI chatbot trained on decades of its own articles*. The Verge. https://www.theverge.com/2024/3/23/24106296/ask-ft-generative-ai-chatbot-answers-tool

Simon, F. M., Nielsen, R. K., & Fletcher, R. (2025). *Generative AI and News Report 2025: How people think about AI's role in journalism and society*. Reuters Institute for the Study of Journalism. https://doi.org/10.60625/RISJ-5BJV-YT69

Simon, F. M., Nielsen, R. K., & Fletcher, R. (2025). *Generative AI and News Report 2025: How people think about AI's role in journalism and society*. Reuters Institute for the Study of Journalism. https://doi.org/10.60625/RISJ-5BJV-YT69

Simon, F. M. (2024). *Artificial Intelligence in the News. How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena* (p. 46). Tow Center for Digital Journalism, Columbia University. https://doi.org/10.7916/ncm5-3v06

Simon, F. M. (2025). Rationalisation of the news: How AI reshapes and retools the gatekeeping processes of news organisations in the United Kingdom, United States and Germany. *New Media & Society*, 14614448251336423. https://doi.org/10.1177/14614448251336423

Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012). *The institutional logics perspective: A new approach to culture, structure, and process*. Oxford University Press.

Toff, B., & Simon, F. M. (2024). "Or They Could Just Not Use It?": The Dilemma of AI Disclosure for Audience Trust in News. *The International Journal of Press/Politics*. https://doi.org/10.1177/19401612241308697

Zoizner, A., Matthes, J., Corbu, N., de Vreese, C. H., Esser, F., Koc-Michalska, K., Schemer, C., Theocharis, Y., & Zilinsky, J. (2025). Can AI-Attributed News Challenge Partisan News Selection? Evidence from a Conjoint Experiment. *The International Journal of Press/Politics*, *0*(0). https://doi.org/10.1177/19401612251342679