Article

The International Journal of Press/Politics
1–28

DOI: 10.1177/19401612261442761
journals.sagepub.com/home/hij
Sage

Going with the Mainstream: Exploring GPT Representation of Journalistic Culture

Taewoo Kang[1], Tim Vos[1], Thomas Hanitzsch[2], Neil Thurman[2], Imke Henkel[3], Sina Thäsler-Kordonouri[2], and Wiebke Loosen[4]

Abstract

As generative AI becomes increasingly integrated into journalism, questions about its implications for journalistic practice grow more urgent. Despite the rise of AI systems, news organizations often lack an informed understanding of how these systems operate, potentially undermining journalists' capacity to exercise agency in AI-assisted news production. This study explores how large language models (LLMs), such as ChatGPT, can reflect journalistic value systems. It does this by prompting the GPT-4o model to respond to survey items from the Worlds of Journalism Study that measure journalistic role perceptions, epistemologies, and ethics. These responses were compared to actual survey data from the US, UK, and Germany. Our findings suggest that GPT-4o's outputs correspond most strongly to the survey responses of politically centrist, full-time journalists whose employers have a TV background, while showing less alignment with right-leaning, part-time, and non-degree-holding journalists. These propensities were most pronounced in the German dataset. While we do not assert that GPT-4o's cultural orientation directly translates to its journalistic applications, the results reveal which groups of journalists' perspectives the LLM is more—or less—likely to reflect. By mapping this alignment, the study offers an empirical point of departure for guiding how journalists can maintain agency

[1]College of Communication Arts and Sciences, Michigan State University, East Lansing, USA
[2]Department of Media and Communication, LMU Munich, München, Germany
[3]Faculty of Arts, Humanities and Cultures, University of Leeds, UK
[4]Leibniz Institute for Media Research, Hans-Bredow-Institut (HBI), Hamburg, Germany

Corresponding Author:
Taewoo Kang, College of Communication Arts and Sciences, Michigan State University, 404 Wilson Road, East Lansing, MI 48824, USA.
Email: kangtaew@msu.edu

in their use of LLMs, and thereby contributes to ongoing discourse about the ethical, epistemological, and institutional dimensions of AI in journalism.

Keywords
large language models, journalistic culture, worlds of journalism study, algorithmic bias, transparency, field theory

Introduction

As attempts to integrate generative AI into newsrooms continue to grow (see Silver 2025, May 7), concerns about its influence on journalism are intensifying (Beckett and Yaseen 2023; Thomson et al. 2025). While the use of generative AI tools has not yet become "taken-for-granted" in journalists' day-to-day work, many journalists report that a limited understanding of how these systems operate is a key risk in journalism's turn toward AI (Minnesota Journalism Center 2025, April 24). The concerns reflect epistemic constraints within the journalistic field, with anxieties about AI's "potential impact on journalistic autonomy, such as algorithmic biases and the loss of editorial control" (Cools and De Vreese 2025: 1) persisting.

These "reasonably foreseeable risks" (Kieslich et al. 2024: 2) have prompted a growing consensus around the need for ethical frameworks tailored to AI's role in journalism. Scholars propose measures that range from the labeling of AI-generated content to more robust editorial guidelines for human–machine collaboration (Paik 2023; Piasecki et al. 2024). However, beyond technical or regulatory fixes, the integration of generative AI, such as large language models (LLMs), raises deeper questions about journalistic culture.

Recent studies suggest that journalists perceive LLMs not only as a tool but as a force that challenges their occupational identity (Van Dalen 2024). In response, journalists' interactions with generative AI often reflect what Wu (2024) calls a "value-motivated use," where adoption is guided by the perceived compatibility of AI outputs with journalistic values. Rather than striving to automate tasks, journalists tend to position AI as an assistant—one whose outputs must be verified, edited, and evaluated (Thäsler-Kordonouri and Koliska 2025; Van Dalen 2024). This reflects what field theory describes as boundary work (Bourdieu 2005): journalists seek to reassert their professional jurisdiction by maintaining control over AI-engaged practices (Carlson and Lewis 2019).

Such so-called "human-in-the-loop" approaches are, however, constrained by the platform nature of AI systems, which yield variable outputs depending on algorithmic architectures typically hidden from end users such as journalists (Ronanki et al. 2024). Although journalists engage in post hoc moderation to adjust AI outputs in line with journalistic standards (Wu 2024), the limited intelligibility of these systems continues to constrain journalists' capacity to exercise power over the trajectories of AI integration (Jones et al. 2022).

To safeguard journalistic agency amid the "institutionalization" of generative AI (Thäsler-Kordonouri and Koliska 2025), journalism must be equipped with domain-specific knowledge of how these systems reflect journalistic culture (Hepp et al. 2023). Such knowledge may enable journalists to evaluate the representational tendencies embedded in algorithmic outputs during their interactions with AI. Although the sources of the sociocultural tendencies embedded in LLMs are inherently difficult to *trace*, users, such as journalists, can *steer* these tendencies based on systematic identification of them (Narayanan and Kapoor 2023; Santurkar et al. 2023); that is, such identification may allow the journalism field to navigate shifting boundaries and develop practices that maximize AI's technological utility while safeguarding journalists' professional legitimacy (Arguedas and Simon 2023).

To this end, we analyzed how GPT-4o reflects the value systems of different journalistic groups. Using survey items from the Worlds of Journalism Study (WJS), we prompted the model to respond to items measuring three central dimensions of journalistic culture: role perceptions, epistemologies, and ethical orientations (Hanitzsch 2007). We then compared these responses with surveys of actual journalists from the US, UK, and Germany. In other words, just as we previously asked journalists, we then asked ChatGPT and compared the answers.

Our findings reveal that responses generated by GPT-4o align[1] more closely with journalists who identify as politically centrist, work full-time, and whose employers have a TV background. By contrast, the model exhibits weaker alignment with right-leaning, part-time, and non-degree-holding journalists. Cross-country comparisons further suggest that GPT-4o's responses are most closely aligned with the response distributions of German journalists overall. These findings suggest that the cultural representativeness of LLMs may not follow a singular tendency but instead varies depending on domain specificity and cultural contexts (Argyle et al. 2023).

However, we do not claim that GPT-4o's representativeness of journalistic culture directly translates into its journalistic applications. Rather, our study aims to identify which journalistic value systems are more—or less—reflected in the model configuration within the context of journalistic culture. By comparing latent patterns that exist between targeted algorithmic outputs and human responses, the study provides initial evidence for advancing the intelligibility of AI within the journalism field (Amigo and Porlezza 2025; Hepp et al. 2023; Jones et al. 2022; Van Dalen 2024).

Theory and Literature Review

This study builds on a theorization of journalistic culture and probes the ability of AI systems to meaningfully represent journalistic culture. The study treats an LLM as a *synthetic journalist* and explores how this synthetic journalist aligns with and within journalism cultures in three countries—the US, UK, and Germany.

Journalistic Cultures across Media Systems

Hanitzsch (2007) theorizes journalistic culture "as a particular set of ideas and practices by which journalists, consciously and unconsciously, legitimate their role in

society and render their work meaningful for themselves and others" (p. 369). He suggests a conceptual structure that incorporates three major areas in which journalism cultures are commonly articulated: journalism's institutional roles; epistemologies; and ethical ideologies. The area of institutional roles relates to the normative and actual functions of journalism in society; epistemologies relate to the accessibility of reality and the nature of acceptable evidence; and ethical ideologies relate to journalists' responses to ethical dilemmas. The theoretical dimensions of journalistic roles include interventionism (active-passive), power distance (adversarial-loyal), and market orientation (consumer-citizen). Journalistic epistemology includes objectivism (correspondence-subjectivity) and empiricism (empirical-analytical) dimensions, and ethical orientations vary along poles of idealism (means-outcomes) and relativism (contextual-universal). Differences among journalists on these dimensions, then, represent differences in normative commitments about journalism's place and practice in society.

Journalism cultures develop within local, regional, and national contexts, where they are responsive to unique social contexts and opportunity structures (Hanitzsch et al. 2011). Even so, evidence suggests journalism cultures around the world bear a resemblance to one another—valuing similar roles and ethical outlooks, for example (Hanitzsch et al. 2019). Variations among journalism cultures, however, are not inconsequential. Neither are variations within journalism cultures, where cross-cutting cleavages can form along demographic and medium differences. Thus, approaches to interventionism, objectivism, and idealism, for example, can and do vary within a journalism culture and across cultures.

Hanitzsch's (2007) conceptualization has been widely applied in numerous empirical studies, serving as the conceptual backbone for the first wave of the WJS (e.g., Hanitzsch et al. 2011; Koroma 2023). This three-dimensional model was also employed in the most recent, third wave of the study, conducted between 2022 and 2024 in seventy-five countries around the globe. This inquiry provides a nuanced perspective on the variations in journalism culture across diverse countries.

Relatedly, the sociology of news (e.g., Shoemaker and Reese 2014; Shoemaker and Vos 2009) points to the ways in which journalism cultures manifest in news content—shaping what becomes news, how news stories are framed, and how prominently news is placed, distributed, and amplified. Different journalistic roles, epistemologies, and ethics lead to different kinds of news content—both within and between countries. For instance, US journalists who embrace a traditional monitorial role, rather than other roles, engage in "practices of relying on elite sources, limiting the diversity of political perspectives, and showing deference and trust for certain authoritative institutions," often to the detriment of citizen and non-elite sources and perspectives (Wolfgang et al. 2021: 1353).

Enter LLMs. Much attention has been focused on their ability—or inability—to shape news content in ways that reflect journalistic judgments that would typically flow from journalists embedded in the dominant journalistic culture—a culture that values autonomy and objectivism (Cools and De Vreese 2025; Wu 2024). What is heretofore largely unknown, however, is whether LLMs can produce an intelligible

picture of journalism. That is, can LLMs describe the nuances of journalism culture—including journalistic roles, epistemologies, and ethics—in ways that comport with journalists' own understandings of journalism culture? Before pursuing this question, we turn to what research to date has been able to tell us about AI as synthetic journalists.

AI Intelligibility in Journalism

We define generative AI as *algorithmic systems that simulate and scale human intellectual capabilities* (Hancock et al. 2020; Mitchell 2019; see also Simon 2024). A prominent example is LLMs, which process natural language stochastically and demonstrate the ability to generate text that is linguistically comparable to human-written content in literacy-driven tasks (Kroon et al. 2024).

Due to these capacities, LLMs are increasingly being deployed in knowledge production domains such as journalism (e.g., Beckett and Yaseen 2023; Thomson et al. 2025). However, the growing exploration of LLMs' journalistic utility raises concerns about potential perils (Cools and Diakopoulos 2024; Møller et al. 2024). Studies highlight multiple discourses within journalism communities about AI's possible risks throughout the journalistic value chain—encompassing news gathering, production, and distribution (BBC 2024; EBU 2024a). A common thread among the discourses is journalists' loss of control and autonomy (Cools and De Vreese 2025; Van Dalen 2024).

As algorithmic systems become a foundational infrastructure of public information technologies (Thurman et al. 2019), platform companies are emerging as central players in the news industry. This shift in status has widened the structural gap between platform companies and news organizations in terms of control over AI systems. In Simon's (2024) interview study of news workers in the US, UK, and Germany, this situation is attributed to the benefits of AI technologies, including "efficiency, ease of use, stability, scalability, and longevity" (p. 159). While these benefits come with the "hidden costs of platform partnerships" (Paik 2023), journalists lack viable alternatives to counter the structural advantages held by platform companies.

Jones et al. (2022) frame this asymmetric power dynamic as *the intelligibility issue of AI in journalism*: that is, "journalists' ability to understand and engage with AI in ways that do not compromise journalistic norms and values" (p. 1731). The limited AI intelligibility restricts journalists in their critical assessments of AI outputs, thereby undermining their ability to control their own AI applications.

Epistemic Challenges for Journalism in the Era of Generative AI

In pursuit of economic benefits, the journalism field continues to explore whether professional tasks—such as news judgment or creative writing—can be automated by LLMs (Thomson et al. 2025). For instance, Diakopoulos (2023, Mar 6) demonstrates that in evaluating the newsworthiness of scientific abstracts published on a pre-print server, ChatGPT exhibited a high correlation with human expert judgments,

suggesting its utility in news discovery. Similarly, Spencer (2025, Mar 20) addresses the capacity of LLMs to unearth overlooked health scandal cases from vast collections of medical documents. Such cases challenge the long-held belief that journalism is insulated from automation and have prompted journalists to state that their "sense of security has disappeared" (Van Dalen 2024: 8).

Developments such as these have led the journalism field to articulate a defensive discourse that questions the feasibility of automating news production (Amigo and Porlezza 2025; Wu 2024). Two primary rationales underlie this view. First, LLMs function as stochastic simulators of human communicative practices rather than autonomous agents (Kok et al. 2021; see Messeri and Crockett 2024), rendering them incapable of fulfilling journalism's normative commitments (Porlezza and Ferri 2022). Second, the presence of latent algorithmic biases and the error of hallucination risk undermining the transparency and factual reliability essential to journalistic integrity (Lin 2022). In turn, these two positions question the legitimacy of the algorithmic systems as journalistic actors (Dörr and Hollnbuchner 2017).

Journalistic Agency and the Human-in-the-Loop Principle

This skepticism about LLMs serves to reaffirm journalists' status as the "story's authoritative agent" (Carlson 2017) while acknowledging the pragmatic imperative to harness the benefits of AI systems (Thäsler-Kordonouri and Koliska 2025). Central to this position is the assertion that human oversight must be preserved within journalist–AI collaboration to uphold the principles of "good journalism" (Wu 2024; see Kovach and Rosentiel 2021). As professional journalism continues to constitute the *gold standard* for news quality, the utility of generative AI outputs is rendered contingent upon human editorial judgment (Ulken 2022).

As such, studies observe that *human-in-the-loop* practices are emerging as ideal principles in the face of the institutionalization of AI journalism. For instance, Wu (2024) describes journalists' use of generative AI as "value-motivated use," a term that encapsulates how the human-oversight principles are performed in the field. This involves three key strategies: (1) refraining from taking AI output at face value, (2) consistently verifying and editing AI-generated content, and (3) thereby upholding core journalistic values. These resonate with what Thäsler-Kordonouri and Koliska (2025) conceptualize as the "hybrid scenario" of AI implementation: human actors and AI systems engage in a collaborative relationship that is marked by mutual influence yet remains under persistent supervision by human judgment.

However, this normative demand for human supervision acknowledges that AI systems can function as a synthetic actor (or a supervisee) within the journalistic field. While AI lacks *autonomous agency*, it comes to share in human agency through interaction (Johnson and Verdicchio 2019; Latour 2005); that is, humans tend to *project* agency onto AI systems, thereby experiencing them as a collective actor. Hepp et al. (2023) explain that this "hybrid agency" gives rise to communicative practices that are distinct from those in newsrooms that have not adopted AI systems (p. 51).

While LLMs do not autonomously possess journalistic sociocultural orientations, collaborating with these models compels journalists to *evaluate* the values projected onto the models' outputs. How such evaluations are carried out, in turn, can shape the dynamics of the news ecosystem (Simon and Isaza-Ibarra 2023). Accordingly, if journalists are not "capable of assuming agency" in evaluating LLM outputs, then it is hard to guarantee the responsible use of journalistic AI, placing structural constraints on journalists' ability to delineate their own professional domain (Helberger et al. 2022: 1615).

Power Dynamics within AI Journalism

If AI journalism is understood as a relationship between human journalists as inherent actors and AI systems as technological actants (Latour 2005), then the social construction of this practice is contingent upon the power dynamics embedded within that relationship (van Rooyen 2013). As Lewis and Westlund (2015) note, the diffusion of power between human and technological agents is manifested in how their respective contributions shape the collaboration's outcome. One expression of these power dynamics is the structural gap between platform companies and news organizations (Jones et al. 2022; Simon 2024). As LLM outputs vary depending on training data, user prompts, and model architecture—which vectorizes the data and user queries in different ways—journalists' limited understanding of such systems further exacerbates the unpredictability of human–AI collaboration (Ronanki et al. 2024).

Research indicates that journalistic AI outputs are neither value-neutral nor guaranteed to be accurate (e.g., Breazu and Katsos 2024; Kuai et al. 2025; Vijay et al. 2024). In turn, if journalists lack *prior* awareness of which voices may be over- or underrepresented at the generative stage, it becomes difficult to assume that *post* moderation alone can comprehensively redress such imbalances (Li and Sinnamon 2024). If journalists lack awareness of AI's relative representativeness, it becomes difficult to claim that they will adequately address the biases allegedly embedded in the AI systems that they seek to control.

Journalists' Normative Understanding of LLMs and AI Practices

Studies report that journalists project both their understanding and personal experiences onto their use of LLMs, particularly the perception that LLM outputs do not reflect journalistic values by default (Thomson et al. 2025). They often presume that generative AI systems are inherently biased and, as such, fail to reflect core journalistic values. The studies suggest that while the cultural outlook embedded in an LLM may not directly translate into its journalistic application, user perception of the model can nonetheless shape the user's communicative practices—that is, how they interact with AI systems (Hepp et al. 2023).

The issue here, however, is that these understandings of AI systems primarily rely on knowledge produced outside the journalism field. Despite discourses on the effective adoption of responsible AI in journalism (e.g., Dodds et al. 2024) and the

formation of AI-related actor networks (e.g., Nordics AI Journalism Network), there is still a need for a systematic establishment of empirical benchmarks within the journalism context; that is, algorithmic biases identified in other domains *should not* be automatically presumed to represent those present in journalistic applications (see Messari and Crockett 2024 for a conceptual review).

Furthermore, when journalists for whom AI have limited intelligibility attempt to exercise oversight over LLM operations, they may have a false sense of how much agency they have, engendered by an "overconfidence in their expectations of the technology" (Van Dalen 2024: 14). At the organizational level as well, when newsroom AI policies lack an evidence-based foundation, they may produce unintended adverse effects by hindering the effective calibration of bias (Bommasani et al. 2025). These possibilities further underscore the need for systematic investigation into LLM algorithmic biases within journalism contexts.

Algorithmic Bias, Fidelity, and Journalistic Value Systems

Bias is broadly defined as "a systematic error in judgments" (Mayson 2018). Operationally, this definition allows social biases in algorithmic outputs to be measured as the degree to which a model's predictions about specific opinions or attitudes diverge from actual targets. In LLMs—and especially in modeling the cognitive and behavioral mechanisms of social groups—bias emerges when systems "are more representative of one class than the other" (Glickman and Sharot 2025: 345). This suggests that algorithmic bias is not limited to overt prejudice (e.g., structural racism) but also encompasses how LLMs' representational stance toward social constructs aligns with or departs from specific group-level opinion distributions.

Argyle et al. (2023) thus define algorithmic bias as "a complex reflection of the many various patterns of association between ideas, attitudes, and contexts present among humans." This definition contrasts with perspectives that treat bias as "a singular, macro-level feature of the model" that generates normatively or ethically wrongful outputs (pp. 337–338). Indeed, a growing body of research demonstrates that LLMs reflect multiple interwoven sociocultural predispositions rather than a single, unified bias (see Kuai et al. 2025).

For an overview, Santurkar et al. (2023) investigate the extent to which LLM responses align with public opinion distributions from Pew Research Center's American Trends Panel (ATP). Their findings indicate that OpenAI's language models tend to more closely mirror the views of liberals (see also Hartmann et al. 2023; Perez et al. 2022). Notably, the level of alignment differs by topic: although liberal perspectives are prevalent, the models sometimes reflect conservative views, particularly on issues such as religion. The studies illustrate that algorithmic bias in LLMs is deeply embedded in sociocultural context and varies depending on the intersection of topical domains and task-specific prompts.

However, computational analyses of how these context-sensitive dynamics manifest within the journalism context require further exploration. Although studies report that LLMs may exhibit cultural variation in how they assess newsworthiness

(Diakopoulos 2023), as well as demonstrating ideological biases in news summarization (Vijay et al. 2024), headline generation (Breazu and Katsos 2024), and fact-checking (Kuznetsova et al. 2025) (see also Bang et al. 2024; Fang et al. 2024; Trhlik and Stenetorp 2024), evaluating such biases in AI *applications* does not equate to assessing a model's *representations* of journalistic value systems.

In particular, the validity of using LLMs to simulate a social group's ideas, attitudes, or behaviors hinges on users' ability to control for the model's algorithmic biases (Argyle et al. 2023). Understanding how an LLM reflects the values of particular journalistic cohorts is therefore pivotal to prompting it in ways that more effectively uphold professional norms.

AI and Journalistic Cultures

The literature on LLMs as synthetic journalists points to several considerations relevant to their ability and desirability to represent journalistic cultures and how they might align with national journalism cultures or cultures built around cross-cutting cleavages. First, the intelligibility issue highlights the centrality of "norms and values" (Jones et al. 2022: 1731) in assessing the efficacy of AI. Second, as stochastic simulators, LLMs present epistemic shortcomings, such that normative commitments may go unrealized (Porlezza and Ferri 2022). Third, the responsible use of LLMs must reflect journalists' awareness of, and control of, their institutional autonomy (Helberger et al. 2022). Fourth, journalists' limited insight into AI's representational tendencies may cause their boundary work (intended to maintain initiative within power dynamics) to falter. Fifth, through the feedback loops of human–AI interaction, a misperception may allow users' own biases to persist while reducing their ability to detect algorithmic biases (Glickman and Sharot 2025). Finally, what remains largely unknown is which journalistic groups' cultural orientations are echoed most (or least) closely by an LLM's representation of journalistic value systems.

In this paper, we operationalize the relative distance between an LLM's representations and those of different journalist groups as the algorithmic bias of the LLM's journalistic culture. Detecting such bias is critical, as recognizing representational tendencies is essential for preserving human agency in interactions with technological actants. The limited awareness of the journalistic cultural orientations that are present in LLMs may restrict journalists' ability to evaluate collaborative outputs in relation to journalistic values: At the input stage, such limitations in AI intelligibility can create uncertainty about how to design prompt engineering or model fine-tuning strategies to mitigate algorithmic bias; at the post hoc editorial stage, they may render the evaluation of outputs overly reliant on journalists' subjective value judgments, thereby hindering the detection of bias (Glickman and Sharot 2025).

This dual constraint, operating both before and after model generation, undermines claims that journalists can secure outcomes aligned with their role perceptions, epistemologies, and ethical orientations. Being able to predict the nature of LLM outputs is, therefore, vital: it enables journalists to assert professional authority and exercise oversight over AI systems in ways that are grounded in their lived experience and

normative commitments. Simply put, such an agency allows journalists to guide the model in a "value-motivated" manner (Wu 2024). While tracing the sources of socio-cultural orientations embedded in LLMs remains inherently challenging, systematic steering becomes more feasible once such model configurations are identified (Narayanan and Kapoor 2023; Santurkar et al. 2023). Accordingly, the identification of these configurations may serve as an empirical benchmark for human actors exercising normative oversight over the models.

Representation of Surveys and Representativeness of LLM Outputs

Just as journalists tend to evaluate AI-generated news content using "upmarket publications as the gold standard" (Van Dalen 2024: 13), social science research that employs LLMs to simulate human cognition often relies on actual human performance—such as expert text annotations—as the *ground truth* (Bail 2024; Karjus 2023; Moser et al. 2024; Ziems et al. 2024). Consistent with this approach, the degree to which AI-generated survey responses resemble those of actual humans often serves as an indicator of an LLM's capacity to model social phenomena within a specific domain (Argyle et al. 2023). The present study adopts such a methodological framework for the first time in the journalism context. Specifically, we prompted an LLM to complete a survey battery derived from the WJS. We then assessed the extent to which the LLM responses aligned with the value orientations reported by actual journalists in the most recent wave of the WJS dataset (see Data and Methods).

Given that journalistic cultures vary both within and across national contexts (Shoemaker and Reese 2014), our analysis operates on two levels: within-country variation—based on demographic, professional, and ideological profiles—and cross-national variation among equivalent subgroups. To structure this inquiry, we adopted a Most Similar Systems Design (Przeworski and Teune 1970) and focused on three advanced democracies: Germany, the United Kingdom, and the United States. Although these countries differ in their media systems (Hallin and Mancini 2004), they share key institutional features, including judicial systems, ideological diversity, public sphere configurations, and the level of economic development. These shared attributes provide a meaningful basis for cross-national and subgroup comparisons.

Moreover, Simon (2024) shows that journalists in these three countries have responded similarly to AI innovations, reflecting institutional isomorphism—the tendency for organizations in a field to converge over time by mimicking successful cases to navigate uncertainty (Caplan and Boyd 2018; Napoli 2014). However, such homogenization may obscure cultural specificities; that is, if algorithmic bias is not well contextualized, convergent AI adoption strategies may risk decontextualizing its journalistic applications—especially where culturally specific assumptions are embedded in practices (Hovy and Prabhumoye 2021; Kroon et al. 2024). These concerns underscore the imperative of assessing LLM cultural representativeness both within and across national contexts. Accordingly, we pose the following research questions:

RQ1. How does an LLM represent different groups of journalists in terms of their journalistic cultures—namely, role perceptions, ethics, and epistemologies?

RQ2. How does the LLM's representation of journalistic cultures—based on the journalist subgroup profiles—vary across the US, UK, and Germany?

RQ3. Which country's journalists—from the US, UK, or Germany—are most closely represented by the LLM in terms of their journalistic cultures?

Data and Methods

Survey Data

Data collection in the three countries followed a common methodological design within the context of the third wave of the WJS (2022–24). The survey was carried out between September 2022 and February 2023 in Germany, between June and November 2023 in the US, and between September and November 2023 in the UK. Samples were of comparable size (Germany: $N = 1{,}221$; UK: $N = 1{,}130$; US: $N = 1{,}326$). Probability samples in the US and UK were drawn from lists of journalists, while the German team extracted a stratified proportional random sample of organizations from which they drew a quota-based random sample of journalists. Data collection was carried out through online surveys in the US and UK; Germany used a mix of telephone and online interviews. Response rates were highest in Germany (15.5%) and somewhat lower in the US (8.2%) and UK (9.0%).

Journalistic Culture Measures

The WJS survey includes blocks of question batteries that cover three dimensions of journalistic culture, all measured on five-point scales. The role perceptions battery consists of twenty-four items that measure how important journalists think it is to fulfill particular roles (such as educating the reader or scrutinizing political leaders). The ethical orientation battery consists of four items that measure journalists' agreement with appropriate responses to ethical dilemmas. The epistemological beliefs battery consists of five global-mandatory and six country-optional questions that measure journalists' levels of agreement with statements about the nature of reality and knowledge. The US and German data include both epistemology batteries (global and optional), while the UK data includes only the mandatory global battery (Table S1).

Adapting the Survey to the LLM Prompt

Using the OpenAI API, we prompted OpenAI's flagship model, *gpt-4o*-2024-*08-06*. A total of thirty-nine survey items were adapted into a standardized prompt format for LLM question-answering tasks. The prompt template replicated the exact original

wording of the WJS survey items, and each question was paired with its corresponding Likert-scale response options (see Hendrycks et al. 2020; Liang et al. 2022 for a methodological overview).[2] To preserve baseline accuracy in journalism contexts (Heseltine and Clemm von Hohenberg 2024), a universal context of WJS was provided once before the start of the template prompting: "Assume you are a journalist, and the survey was conducted in 2023." No additional prompt was engineered that might substantially shift the context relative to human news workers' responses in WJS surveys (Santurkar et al. 2023; see also Lee et al. 2024; Törnberg 2023).[3]

Furthermore, we conducted a series of additional simulations—varying the system prompt to include generic journalist roles, specific national contexts (US, UK, Germany), and no-context baselines—to confirm whether the findings were substantially influenced by the specific wording of the universal context. These tests demonstrate that the model's representational patterns remain consistent across minor prompt variations and are not merely artifacts of specific contextual cues. This step ensures that the simulation maintains high fidelity to the intended setting while remaining robust to the nuances of prompt engineering.[4]

The *temperature* parameter, which determines the model's randomness or creativity, was set to 0.7 in accordance with prior research (e.g., Argyle et al. 2023; Lee et al. 2024). To handle transient errors, the *max retries* parameter was set to its default value of 3. Lastly, following Heseltine and Clemm von Hohenberg's (2024) recommendation, we repeated the prompts to assess the inter-rater reliability of the model's outputs (Krippendorff's $\alpha = .92$).

Evaluating LLM Representativeness

Following Santurkar et al. (2023), the degree to which LLM responses represented the actual survey responses of a group of journalists was assessed using the inverted Wasserstein distance (IWD). The Wasserstein distance offers a measure of how closely the model's prediction aligns with the broader pattern of human responses, allowing for a nuanced understanding of how *off* the prediction is in a distributional sense (Peyré and Cuturi 2019). The IWD provides an optimal value—ranging from 0 to 1—that is comparable across different distributions, quantifying the *cost* of transforming one distribution into another (Villani 2009).

We first computed the Wasserstein distance between the GPT response and the response distribution of each subgroup for each question. The WJS's original profiling items used to compute the subgroup distributions included political view, gender, employment status, main employer's background, and highest education level (see Table S2 for measurement details). Additionally, individual datasets for the US, UK, and Germany were each profiled as separate subsamples, and the response distributions for each country were computed accordingly.

Subgroups with sample sizes of thirty or fewer were excluded from statistical comparisons. For cross-national comparisons, the following groups were excluded: "Other" gender, "Other" employment status, "Telecommunications" or "Other" backgrounds, "Not completed high school," "Completed high school," and "Doctorate." In within-country comparisons, exclusion groups varied by each nation's sample

configuration. The "Other" gender category was excluded from the analyses for the UK and Germany, but was retained for the US ($N=35$). Similarly, the "Completed high school" group was included in analyses for the UK ($N=56$) and Germany ($N=238$), while the doctorate group was included only in Germany's analysis ($N=37$).

We then operationalized *representativeness* by subtracting the normalized Wasserstein distance from 1 (i.e., IWD). For instance, a larger IWD for an item indicates a smaller gap between the LLM response and the central tendency of a subgroup's response distribution, increasing the representativeness.

Results

We began by analyzing the representativeness of GPT-4o's responses across all thirty-nine journalistic items. Next, we conducted a more detailed analysis of the ten items that showed the highest variabilities in the Wasserstein distance between the subgroup response distributions; the larger variability indicates a more pronounced algorithmic bias toward specific subgroups.[5] Depending on the scope of analysis, the subgroups were extracted from either the combined sample of the three countries or individual country samples. All code for the analysis is available (Supplemental Materials).

RQ1: How Does GPT-4o Represent Different Groups of Journalists in Terms of Their Journalistic Culture?

Figure 1 illustrates substantial variation in the model's representativeness across different subgroups. In the compound analysis across all three dimensions, GPT-4o's responses aligned closely with the distributions of groups identifying as centrist, full-time contract workers, employed by outlets with a TV background, and individuals holding a BA degree. In contrast, the model showed weaker alignment with responses from right-leaning individuals, part-time contract workers, those whose main employer has a news agency background, and individuals without a formal degree. No notable gender differences were observed.

A more detailed analysis of each dimension revealed additional differences. With regard to epistemological beliefs, GPT-4o's responses most closely aligned with the distributions of journalists identifying as left-leaning, those with a female gender identity, and those whose main employer had an internet-native background. In the ethical orientation battery, the model's centrist bias became more pronounced, and its responses aligned more closely with those of part-time workers than those of full-time workers. In the role perception dimension, the results were broadly consistent with those observed in the compound analysis (Figure 2).

Subsequently, we examined the Wasserstein distances across all items to identify the ten items with the highest coefficient of variation in representativeness between subgroups (Figure 3). The item with the highest variability was role_Q ("Convey a positive image of political leaders"). GPT-4o's responses were mostly aligned with those of left-leaning and internet-native background groups, while showing greater distance from those of right-leaning and no formal degree groups. The second-highest

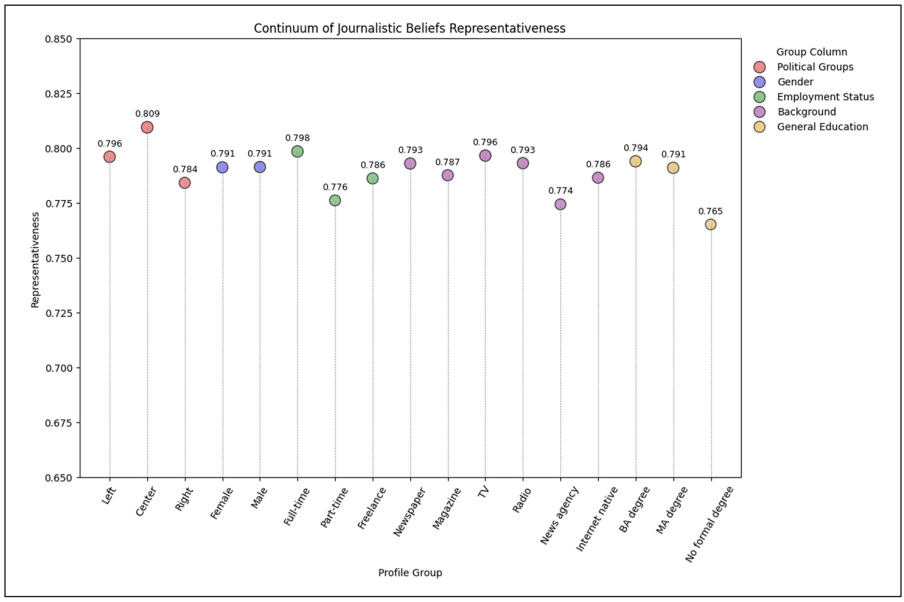


Figure 1. GPT-4o representativeness for subgroups within the combined dataset.

variability was observed for epist1_D ("Things are either true or false, there is no in-between"), where GPT-4o's responses aligned most closely with the centrist subgroup's distributions.

RQ2: How Does GPT-4o's Representativeness Vary Across the US, UK, and Germany?

We conducted country-specific analyses and identified unique patterns that were not apparent in the combined dataset (Tables S4 and S5). With the US sample, a three-dimensional compound analysis revealed that GPT-4o responses closely aligned with male journalists but poorly aligned with journalists identifying as "Other" gender. Unlike in the combined dataset, GPT-4o's representativeness was higher for responses from journalists whose main employer had a news agency background than for those whose main employer had a magazine background. The centrist bias persisted in the dimension compound analysis, but for epistemological beliefs and ethical orientations, GPT-4o's representativeness was higher for left-leaning groups than for centrist or right-leaning groups. In addition, variability in representativeness between groups was most pronounced for role_Q ("Convey a positive image of political leaders"; CV = 0.375). Notably, the top nine items with the highest variability were from the role perception battery.

With the UK sample, the compound analysis showed a pronounced left-leaning bias (IWD for left-leaning: 0.80, centrist: 0.77, right-leaning: 0.75) across all three

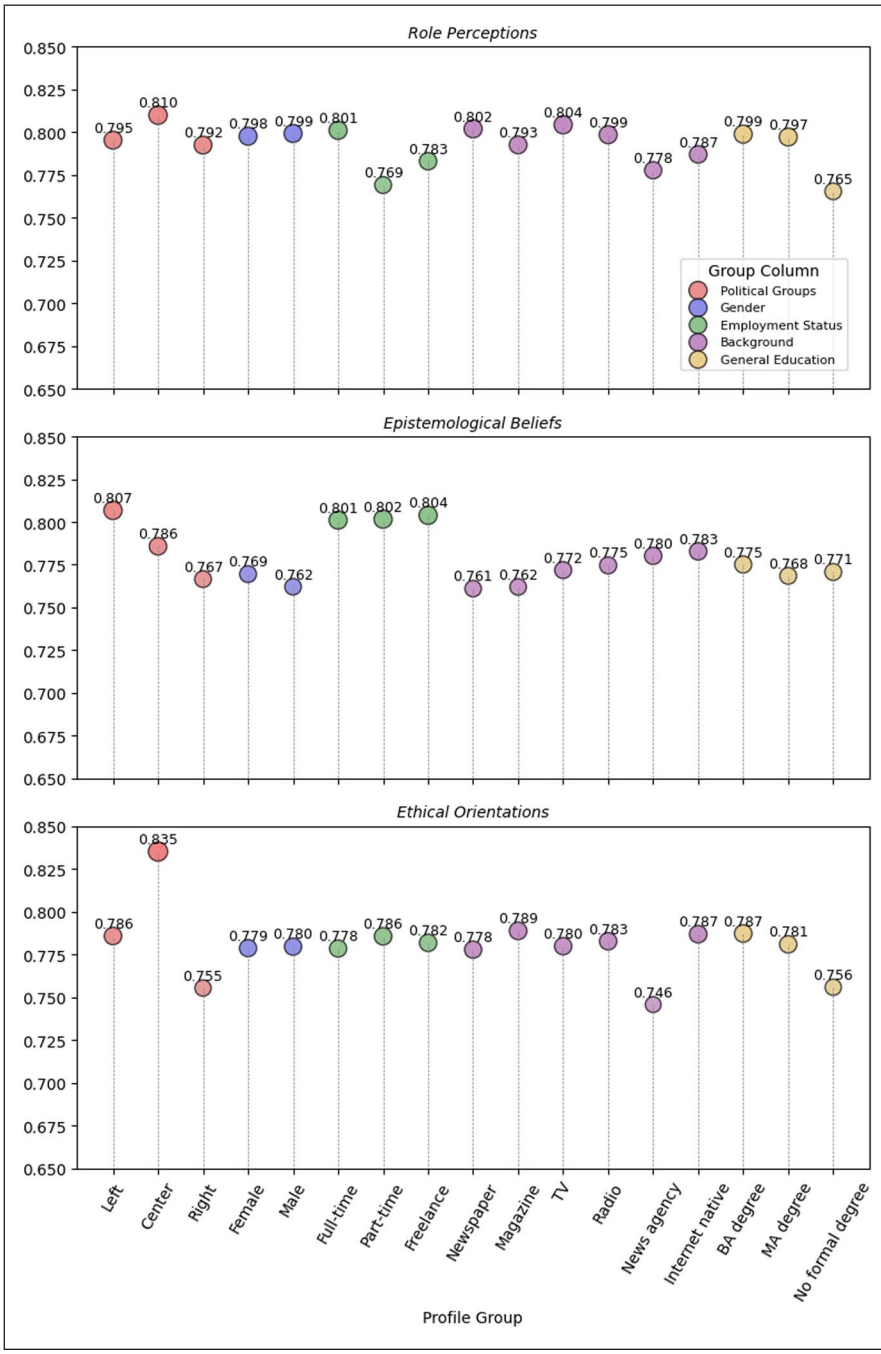
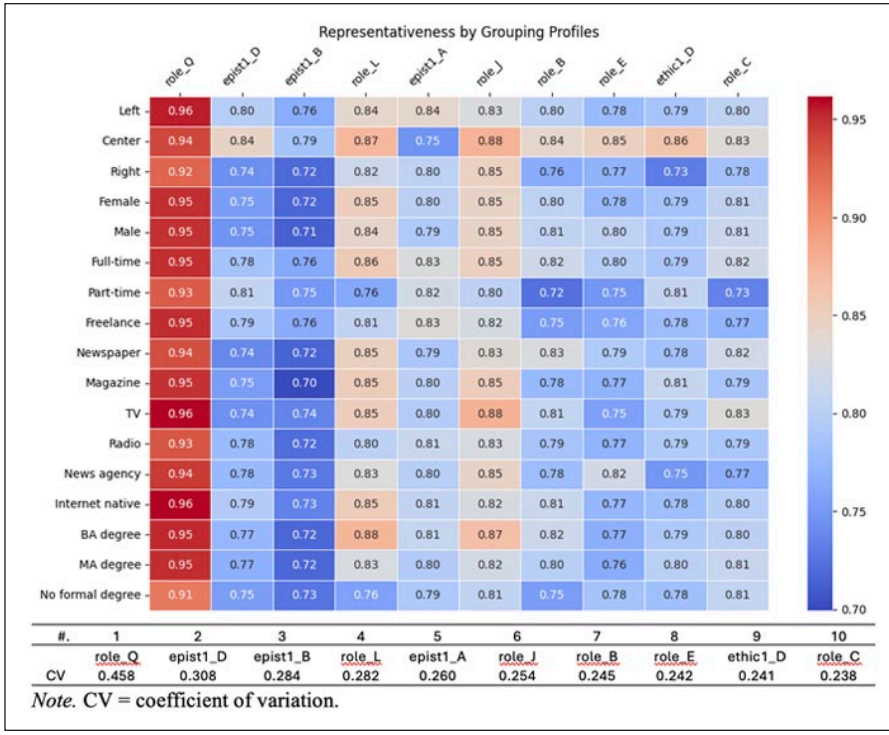


Figure 2. GPT-4o representativeness across belief dimensions within the combined dataset.



#.	1	2	3	4	5	6	7	8	9	10
	role_Q	epist1_D	epist1_B	role_L	epist1_A	role_J	role_B	role_E	ethic1_D	role_C
CV	0.458	0.308	0.284	0.282	0.260	0.254	0.245	0.242	0.241	0.238

Note. CV = coefficient of variation.

Figure 3. Top 10 high variability item comparisons within the combined dataset.
Note. CV = coefficient of variation.

batteries. In particular, high representativeness for the TV-background group was more evident than in the combined dataset. With regard to variability analysis, role_Q exhibited the highest variability in the UK sample, just as it did in the US sample. Another noteworthy item was role_P ("Support government policy"), which had the second-highest variability. GPT-4o's responses aligned least with the distribution of right-leaning groups.

The analysis of the German sample showed the strongest centrist bias in GPT-4o's responses among the three countries: In the ethics dimension, the representational gap between centrist and right-leaning groups exceeded 10.8% in IWD. Moreover, GPT representativeness for the TV-background group in Germany was less pronounced than in the UK and the US, while response distributions for MA degree holders showed notable alignment. In contrast to the three-country combined or US and UK analyses, the item with the highest variability for the German dataset was epist_D ("Things are either true or false, there is no in-between"). For this item, GPT-4o's responses aligned least with the right-leaning group (IWD for left-leaning: 0.91, centrist: 0.93, right-leaning: 0.80). That is, GPT representation of epistemological perception on dichotomous fact-judgments was closest to that of left-leaning journalists in Germany.

RQ3: Which Country's Journalistic Cultures are Most Closely Represented by GPT-4o?

After conducting country-specific analyses, we returned to the combined dataset for a cross-country comparison. Using the *country* code as a grouping factor, we computed the IWDs to assess the alignment between GPT-4o responses and the country-aggregated response distributions.

As illustrated in Figure 4, GPT-4o's responses showed the closest alignment with German journalists across dimensions. While the German sample exhibited narrow differences in IWDs when compared with the UK in the epistemology dimension and the US in the role perception dimension, it demonstrated substantially higher GPT-4o representativeness than both the UK and US data in the ethics dimension. Collectively, the three-dimensional compound analysis indicated that GPT-4o's responses to the journalistic culture battery aligned most closely with the response distributions of German journalists (bottom pane).

Discussion

Recent scholarship has shown that journalists are constructing discourses to reassert their professional authority amid the rise of generative AI (Amigo and Porlezza 2025; Van Dalen 2024). Central to these discourses are concerns that, while generative AI might ostensibly perform conventional journalistic tasks, its outputs lack core journalistic values (Porlezza and Ferri 2022; Thomson et al. 2025). These concerns have contributed to the growing emphasis on *human-in-the-loop* practices, which assert the necessity of human oversight over AI-generated content (Thäsler-Kordonouri and Koliska 2025; Wu 2024).

However, calls for human oversight also expose the fact that AI systems have limited intelligibility for journalists (Jones et al. 2022; Simon 2024); that is, journalistic engagement with AI often remains constrained to practices that rely on post hoc edits or prompts that fail to sufficiently reflect journalistic contexts (Nishal and Diakopoulos 2024). Limited AI intelligibility, in turn, may constrain journalists' capacity to critically assess LLM outputs. Such structural constraints on oversight capacities raise the possibility that if the institutionalization of AI journalism unfolds in a way that sees technological actants primarily shape initial outputs while human agents are relegated mainly to ex-post oversight roles, then asymmetries in human–AI power dynamics could persist (Latour 2005; Lewis and Westlund 2015). These power asymmetries can be particularly problematic, given it is still the journalist who bears the editorial responsibility for content produced with the support of AI tools.

This study addresses the challenge posed by these asymmetries by offering insight into how GPT-4o reflects journalistic culture. Using LLM simulations based on survey questions asked in the WJS, we measured GPT-4o's alignment with journalist subgroups across three dimensions: role perceptions, epistemological beliefs, and ethical orientations (Hanitzsch 2007). Our analysis revealed consistent alignment with politically centrist, full-time, and TV-background journalists, with the strongest overall alignment found in the German dataset.

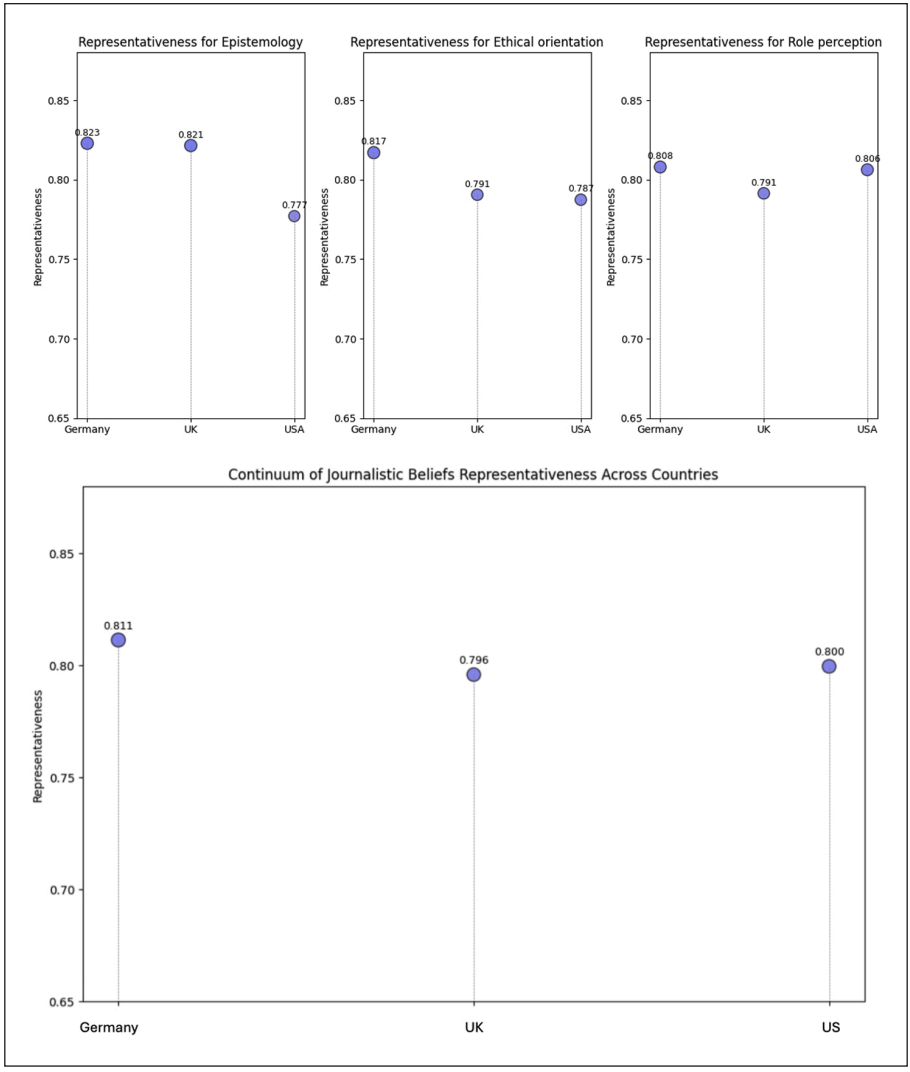


Figure 4. GPT-4o representativeness across belief dimensions in the three countries.

These patterns suggest that GPT-4o does not represent a singular cultural orientation but rather reflects intertwined configurations of journalistic values in various contexts (Argyle et al. 2023). This highlights the need for localized strategies in practicing value-motivated uses of AI (Wu 2024). While LLMs tend to *go with the mainstream* (Bender et al. 2021), the groups associated with those major views can vary across societies. Our findings show that centrist groups in Germany culturally align more closely with left-leaning groups in the US and UK. That is, a society's journalistic cultural landscape must be considered when advocating for the integration of diverse

perspectives into LLM uses, as such advocacy depends on which groups are positioned as "mainstream" within that specific media system. Despite the growing institutional isomorphism in journalists' responses to AI across Germany, the UK, and the US (Simon 2024), this study illustrates that cultural nuances remain vital for assessing LLMs' journalistic feasibility.

As such, these findings offer empirical evidence for concerns about potential algorithmic biases embedded in AI—within journalistic contexts (Cools and De Vreese 2025). Specifically, GPT-4o's representational tendencies were in line with ideological skews observed not only in other domains but also in several journalistic applications (e.g., Bang et al. 2024; Fang et al. 2024; Trhlik and Stenetorp 2024). Moreover, the model's closer alignment with majority groups—such as television journalists and full-time workers—points to the need for continued vigilance among human journalists regarding the potential systematic omission of minority perspectives in LLM outputs.

Accordingly, this study emphasizes the importance of the journalism field recognizing the cultural configurations embedded in LLMs as a means of ensuring ethical AI practices (Arguedas and Simon 2023). At the same time, it provides a rationale for human-in-the-loop approaches—from prompt design to post-editing.

The study also yields evidence-based policy implications (Bommasani et al. 2025). News organizations and journalism education institutions should incorporate AI literacy—particularly with respect to the mainstream-driven orientations of LLMs—within both organizational policy frameworks and the design of training programs (Diakopoulos and Koliska 2017). Understanding which journalistic cultures are more—or less—reflected in LLMs and how such algorithmic biases can be steered (Narayanan and Kapoor 2023) should be the basis for editorial guidelines and ethical standards.

Limitations and Conclusion

Several limitations must be acknowledged. First, the WJS data are based on quota or probability samples rather than on a full census. Although they offer rigorous approximations of national journalist populations (Wahl-Jorgensen and Hanitzsch 2019), the generalizability of subgroup-level findings should be interpreted with care due to small sample sizes for some groups (e.g., regarding education and employer background). Second, the analysis was based on an explorative implementation of a specific configuration of GPT-4o. Given that LLM output is contingent on fine-tuning and training data (Ouyang et al. 2022), alignment levels with journalistic cultures may vary between model versions and types (Buyl et al. 2024). This calls for longitudinal tracking through repeated benchmarking (Messeri and Crockett 2024).

Relatedly, it is necessary to articulate the scholarly position of the analysis, given the inherent variability of LLM architectures and their outputs (Barrie et al. 2024). We recognize that the outputs of our LLM-driven survey simulation are subject to variation due to multiple technical factors (e.g., prompt engineering, parameter settings, model specification). Although significant inter-coder reliability between the outputs was observed across five additional simulation variants (see Notes 3 and 4), this does not indicate that variability in model reproduction can be fully controlled.

However, it is equally important to recall that the validation of LLM-based research tools remains unconsolidated within broader social science scholarship (Peng and Yang 2025). In text annotation, for example, researchers have experimented with Monte Carlo techniques (Tolochko et al. 2025), averaging multiple outputs (Zhao et al. 2025), and leveraging double execution (Heseltine and Clemm von Hohenberg 2024) to validate the LLM coding outputs. Each of these approaches has distinct strengths and limitations, yet the field has not reached a consensus on establishing a common standard. Adjudicating among them lies beyond the scope of this study.

That said, this indeterminacy does not constitute a reason to avoid the scholarly use of LLMs in addressing our research questions. Rather, we contend that the value of this study lies in serving as a *catalyst* for academic discourse on algorithmic bias in journalistic applications. That is, we position our study as a demonstrative application of algorithmic bias identification rather than a confirmatory test. Our LLM survey simulation was deliberately designed to emulate journalist groups at large. The objective herein is not to consolidate GPT responses with a fixed standard, but to observe potential patterns in which journalist groups' perspectives align more—or less—closely with model outputs, under hypothetical conditions of integration into journalistic AI practice. The present analysis thus furnishes such "what if" scenarios and, in doing so, supports counterfactual reasoning about journalistic collaboration with LLMs.

Aligning with our position, scholars increasingly contend that in the scientific use of LLMs, these models should be viewed not only as methodological tools but also as conceptual resources for abductive theorization (Bail 2024). This perspective resonates with Margolin's (2019) argument that computational social science ought to privilege theoretical plausibility over generalizability to generate novel insights into established research programs, while strategically pursuing targeted analyses (see also Davidson and Karell 2025).

Likewise, the purpose of our analysis is not to produce universally replicable simulations but to demonstrate possible model patterns when models are applied to inquiries into journalistic culture. In doing so, the study highlights potential group-specific interactions that might emerge if LLMs were incorporated into newsroom practice. The theoretical proposition that we advance—that variation in model–user alignment may shape divergent news production outcomes—offers a basis for subsequent empirical testing within established research traditions.

One promising future avenue would be to test whether journalists' prompting strategies change when they are informed about a model's representational tendencies. Researchers could then measure how this awareness affects both prompting and post-editing practices. Another extension involves panel designs using multiple LLMs (e.g., Mistral, based in Europe, and DeepSeek, based in China) to compare cross-model representations. This is all the more important given that journalistic practice increasingly involves multiple model comparisons (EBU 2024b). Combined with WJS data, such an approach could become a powerful benchmarking tool for journalism studies (see Buyl et al. 2024; Santurkar et al. 2023).

Ultimately, this study highlights the fact that effective AI integration in journalism requires more than technical safeguards. It demands a social understanding of how

journalistic cultures and practices are encoded, reproduced, and disrupted by generative systems. In this regard, it is worth referencing Messeri and Crockett (2024), who emphasize the need for cognitive and demographic diversity among those evaluating AI outputs, and warn against a monoculture that privileges a singular (major/mainstream) perspective.

This principle is equally vital to journalism, which should serve as a pluralistic lens in democratic societies (Porto 2007). Therefore, assessment of AI's role in journalism must move beyond the balancing of efficiency against risk. It requires a critical rethinking of how journalistic authority and cultural plurality can be sustained amid algorithmic shifts. Key to this effort is evaluating the alignment between the perspectives that newsrooms aim to present and those that AI systems are likely to *re*present. By mapping these socio-technical configurations, this study aims to inform journalistic communities as they navigate and shape the evolving landscape of AI-driven news ecosystems.

ORCID iDs

Taewoo Kang https://orcid.org/0000-0002-4995-5145
Tim Vos https://orcid.org/0000-0002-7174-5574
Thomas Hanitzsch https://orcid.org/0000-0002-7104-6300
Neil Thurman https://orcid.org/0000-0003-3909-9565
Imke Henkel https://orcid.org/0000-0001-7754-7977
Sina Thäsler-Kordonouri https://orcid.org/0000-0003-3455-9315
Wiebke Loosen https://orcid.org/0000-0002-2211-2260

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Throughout this paper, the terms "align" or "alignment" refer specifically to the degree to which LLM survey simulations resemble the response distributions of given human participant groups. Specifying this terminological usage is essential given that in computer science—where the term "alignment" originates—the terms carry multiple connotations, ranging from output personalization to prompt responsiveness to simulation accuracy.
2. A common concern with the QA-template approach is that a language model may retrieve memorized information rather than process queries stochastically. This risk is minimal here. The WJS third-wave data remains under embargo and is not publicly available, making it highly unlikely that GPT-4o was trained on it. Even if earlier WJS waves or related studies had been included in training, any resemblance in output cannot be attributed directly to

such exposure. Our templates excluded WJS-specific keywords, further reducing this possibility. In short, the study relied on non-public data and an established API-based method designed to avoid methodological circularity and to elicit the model's algorithmic propensities rather than memory retrieval.

3. All prompt engineering was conducted in English, with no country-specific information included during response generation or analysis. This reflects English's role as the "[g]lobal lingua franca, the default language of the [GPT-4] model" (Kuai et al. 2025: 9), and the model's original training on corpora rooted in Anglophone socio-cultural contexts (Buyl et al. 2024). The study seeks to identify the journalistic cultural orientations represented by GPT-4o. Introducing other languages or country codes during prompting might obscure this baseline in the journalism contexts by adding additional linguistic variables. Nevertheless, such inputs are essential for assessing the model's capacity to simulate journalistic values specific to different cultural or national contexts—particularly those expressed in non-English languages (Santurkar et al. 2023). In this study, German serves as such a case. We therefore conducted a supplementary survey simulation using German-language prompts and performed a comparative analysis with the WJS German data. The results based on German prompting were overall highly similar to those from English prompting (Krippendorff's $\alpha = .85$). The pattern of GPT representativeness for German journalists likewise aligned with the main analysis, although subtle quantitative differences were observed for certain items. Detailed results are reported in the Supplemental Materials (S6).
4. Detailed descriptions of these five system prompt variations, the results of the additional simulations, and the corresponding inter-coder reliability scores are provided in Supplemental Materials (S7 and S8).
5. Variability was calculated using the coefficient of variation (CV), defined as the standard deviation divided by the mean, within the normalized Wasserstein metric (Reed et al. 2002).

References

Amigo, L., and C. Porlezza. 2025. "Journalism will Always Need Journalists: The Perceived Impact of AI on Journalism Authority in Switzerland." *Journalism Practice*. Published ahead of print April 10. https://doi.org/10.1080/17512786.2025.2487534.

Arguedas, A. R., and F. M. Simon. 2023. *Automating Democracy: Generative AI, Journalism, and the Future of Democracy*. Balliol Interdisciplinary Institute, University of Oxford.

Argyle, L. P., E. C. Busby, N. Fulda, J. R. Gubler, C. Rytting, and D. Wingate. 2023. "Out of One, Many: Using Language Models to Simulate Human Samples." *Political Analysis* 31, no. 3: 337–51. https://doi.org/10.1017/pan.2023.2.

Bail, C. A. 2024. "Can Generative AI Improve Social Science?." *Proceedings of the National Academy of Sciences* 121, no. 21: e2314021121. https://doi.org/10.1073/pnas.2314021121.

Bang, Y., D. Chen, N. Lee, and P. Fung. 2024. "Measuring Political Bias in Large Language Models: What Is Said and How It Is Said." *arXiv preprint arXiv:2403.18932*. https://doi.org/10.48550/arXiv.2403.18932.

Barrie, C., A. Palmer, and A. Spirling. 2024. "Replication for Language Models Problems, Principles, and Best Practice for Political Science." http://arthurspirling.org/documents/BarriePalmerSpirling_TrustMeBro.pdf.

BBC. 2024. "Embedding the Audience: Putting Audiences at the Heart of Generative AI." https://www.bbc.co.uk/aboutthebbc/documents/what-do-people-think-of-generative-ai.pdf

Beckett, C., and M. Yaseen. 2023. "Generating Change. A global Survey of What News Organisations are Doing with AI. The London School of Economics and Political Science." https://www.aiunplugged.io/wp-content/uploads/2023/10/Generating-Change-A-global-survey-of-what-news-organisations-are-doing-with-AI-By-Cyber-Gear.pdf.

Bender, E. M., T. Gebru, A. McMillan-Major, and S. Shmitchell. 2021, March. "On the Dangers of Stochastic Parrots: Can Language Models be Too big?." *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–23). https://doi.org/10.1145/3442188.3445922.

Bommasani, R., S. Arora, J. Chayes, et al. 2025. "Advancing Science-and Evidence-Based AI Policy." *Science* 389, no. 6759: 459–61. https://doi.org/10.1126/science.adu8449.

Bourdieu, P. 2005. "The Political Field, the Social Science Field, and the Journalistic Field." In *Bourdieu and the Journalistic Field* (pp. 29–47), edited by R. Benson and E. Neveu. Polity.

Breazu, P., and N. Katsos. 2024. "ChatGPT-4 as a Journalist: Whose Perspectives is it Reproducing?" *Discourse and Society*, 35, no. 6: 687–707. https://doi.org/10.1177/09579265241251479.

Buyl, M., A. Rogiers, S. Noels, et al. 2024. "Large Language Models Reflect the Ideology of Their Creators." *arXiv preprint arXiv:2410.18417*. https://doi.org/10.48550/arXiv.2410.18417

Caplan, R., and D. Boyd. 2018. "Isomorphism Through Algorithms: Institutional Dependencies in the Case of Facebook." *Big Data & Society* 5, no. 1: 1–12. https://doi.org/10.1177/2053951718757253.

Carlson, M. 2017. *Journalistic Authority: Legitimating News in the Digital Era*. Columbia University Press.

Carlson, M., and S. C. Lewis. 2019. "Boundary Work." In *The Handbook of Journalism Studies* (pp. 123–35), edited by K. Wahl-Jorgensen and T. Hanitzsch. Taylor & Francis.

Cools, H., and C. H. De Vreese. 2025. From Automation to Transformation with AI-Tools: Exploring the Professional Norms and the Perceptions of Responsible AI in a News Organization. *Digital Journalism*. Published ahead of print May 17. https://doi.org/10.1080/21670811.2025.2505982.

Cools, H., and N. Diakopoulos. 2024. "Uses of Generative AI in the Newsroom: Mapping Journalists' Perceptions of Perils and Possibilities." *Journalism Practice*. Published ahead of print August 26. https://doi.org/10.1080/17512786.2024.2394558

Davidson, T., and D. Karell. 2025. "Integrating Generative Artificial Intelligence into Social Science Research: Measurement, Prompting, and Simulation." *Sociological Methods & Research*. Published ahead of print August 1. https://doi.org/10.1177/00491241251339184.

Diakopoulos, N. 2023, March 6. "Finding Newsworthy Documents Using Generative AI." https://generative-ai-newsroom.com/finding-newsworthy-documents-using-generative-ai-a43a41a90e30.

Diakopoulos, N., and M. Koliska. 2017. "Algorithmic Transparency in the News Media." *Digital Journalism* 5, no. 7: 809–28. https://doi.org/10.1080/21670811.2016.1208053.

Dodds, T., A. Vandendaele, F. M. Simon, et al. 2024. "The Impact of Knowledge Silos on Responsible AI Practices in Journalism." *arXiv preprint arXiv:2410.01138*. https://doi.org/10.48550/arXiv.2410.01138

Dörr, K. N., and K. Hollnbuchner. 2017. "Ethical Challenges of Algorithmic Journalism." *Digital Journalism* 5, no. 4: 404–19. https://doi.org/10.1080/21670811.2016.1167612.

EBU. 2024a. "News Report 2024: Trusted Journalism in the Age of Generative AI." https://www.ebu.ch/guides/open/report/news-report-2024-trusted-journalism-in-the-age-of-generative-ai.

EBU. 2024b. "AiDitor: Pioneering AI-Driven Innovation at ORF." https://tech.ebu.ch/news/2024/08/aiditor-pioneering-ai-driven-innovation-at-orf

Fang, X., S. Che, M. Mao, et al. 2024. "Bias of AI-generated Content: An Examination of News Produced by Large Language Models." *Scientific Reports* 14, no. 1: 5224. https://doi.org/10.1038/s41598-024-55686-2.

Glickman, M., and T. Sharot. 2025. "How human–AI feedback loops alter human perceptual, emotional and social judgements." *Nature Human Behaviour* 9, no. 2: 345–59. https://doi.org/10.1038/s41562-024-02077-2.

Hallin, D. C., and P. Mancini. 2004. *Comparing Media Systems: Three Models of Media and Politics*. Cambridge University Press.

Hancock, J. T., M. Naaman, and K. Levy. 2020. "AI-mediated Communication: Definition, Research Agenda, and Ethical Considerations." *Journal of Computer-Mediated Communication* 25, no. 1: 89–100. https://doi.org/10.1093/jcmc/zmz022.

Hanitzsch, T. 2007. "Deconstructing Journalism Culture: Toward a Universal Theory." *Communication Theory* 17, no. 4: 367–85. https://doi.org/10.1111/j.1468-2885.2007.00303.x.

Hanitzsch, T., F. Hanusch, C. Mellado, et al. 2011. "Mapping Journalism Cultures Across Nations: A Comparative Study of 18 countries." *Journalism studies* 12, no. 3: 273–93. https://doi.org/10.1080/1461670X.2010.512502.

Hanitzsch, T., F. Hanusch, J. Ramaprasad, and A. De Beer. 2019. *Worlds of Journalism: Journalistic Cultures Around the Globe*. Columbia University Press.

Hartmann, J., J. Schwenzow, and M. Witte. 2023. "The Political Ideology of Conversational AI: Converging Evidence on ChatGPT's Pro-environmental, Left-libertarian Orientation." *arXiv preprint arXiv:2301.01768*. https://doi.org/10.48550/arXiv.2301.01768.

Helberger, N., M. van Drunen, J. Moeller, S. Vrijenhoek, and S. Eskens. 2022. "Towards a Normative Perspective on Journalistic AI." *Digital Journalism* 10, no. 10: 1605–1626. https://doi.org/10.1080/ 21670811.2022.2152195.

Hendrycks, D., C. Burns, S. Basart, et al. 2020. "Measuring Massive Multitask Language Understanding." *arXiv preprint arXiv:2009.03300*. https://doi.org/10.48550/arXiv.2009.03300.

Hepp, A., W. Loosen, S. Dreyer, et al. 2023. "ChatGPT, LaMDA, and the hype around communicative AI: The automation of communication as a field of research in media and communication studies." *Human-Machine Communication* 6, no. 1: 4. https://doi.org/10.30658/hmc.6.4.

Heseltine, M., and B. Clemm von Hohenberg. 2024. "Large Language Models as a Substitute for Human Experts in Annotating Political Text." *Research & Politics* 11, no. 1: 20531680241236239. https://doi.org/10.1177/20531680241236239.

Hovy, D., and S. Prabhumoye. 2021. "Five Sources of Bias in Natural Language Processing." *Language and linguistics Compass* 15, no. 8: e12432. https://doi.org/10.1111/lnc3.12432.

Johnson, D. G., and M. Verdicchio. 2019. "AI, Agency and Responsibility: The VW Fraud Case and Beyond." *AI & Society* 34: 639–47. https://doi.org/10.1007/s00146-017-0781-9.

Jones, B., R. Jones, and E. Luger. 2022. "AI 'Everywhere and Nowhere': Addressing the AI Intelligibility Problem in Public Service Journalism." *Digital Journalism* 10, no. 10: 1731–55. https://doi.org/10.1080/21670811.2022.2145328.

Karjus, A. 2023. "Machine-assisted Mixed Methods: Augmenting Humanities and Social Sciences with Artificial Intelligence." *arXiv preprint arXiv:2309.14379*. https://doi.org/10.48550/arXiv.2309.14379

Kieslich, K., N. Diakopoulos, and N. Helberger. 2024. "Anticipating Impacts: Using Large-scale Scenario-writing to Explore Diverse Implications of Generative AI in the News Environment." *AI and Ethics*, 1–23. https://doi.org/10.1007/s43681-024-00497-4.

Kok, K. P. W., A. M. C. Loeber, and J. Grin. 2021. "Politics of Complexity: Conceptualizing Agency, Power and Powering in the Transitional Dynamics of Complex Adaptive Systems." *Research Policy* 50: 104183–11. https://doi.org/10.1016/j.respol.2020.104183.

Koroma, S. B. 2023. *Journalism Cultures in Sierra Leone: Between Global Norms and Local Pressures*. Palgrave Macmillan.

Kovach, B., and T. Rosenstiel. 2021. *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect* (4th ed.). Penguin Random House.

Kroon, A., K. Welbers, D. Trilling, and W. van Atteveldt. 2024. "Advancing Automated Content Analysis for a New Era of Media Effects Research: The Key Role of Transfer Learning." *Communication Methods and Measures* 18, no. 2: 142–62. https://doi.org/10.1080/19312458.2023.2261372.

Kuai, J., C. Brantner, M. Karlsson, E. Van Couvering, and S. Romano. 2025. "AI Chatbot Accountability in the Age of Algorithmic Gatekeeping: Comparing Generative Search Engine Political Information Retrieval Across Five Languages." *New Media & Society*. https://doi.org/10.1177/14614448251321162

Kuznetsova, E., M. Makhortykh, V. Vziatysheva, M. Stolze, A. Baghumyan, and A. Urman. 2025. "In Generative AI We Trust: Can Chatbots Effectively Verify Political Information? *Journal of Computational Social Science* 8, no. 1: 15. https://doi.org/10.1007/s42001-024-00338-8.

Latour, B. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Lee, S., T. Q. Peng, M. H. Goldberg, et al. 2024. "Can Large LANGUAGE models Estimate Public Opinion About Global Warming? An Empirical Assessment of Algorithmic Fidelity and Bias." *PLOS Climate* 3, no. 8: e0000429. https://doi.org/10.1371/journal.pclm.0000429.

Lewis, S. C., and O. Westlund. 2015. "Actors, Actants, Audiences, and Activities in Cross-media News Work: A Matrix and a Research Agenda." *Digital Journalism* 3, no. 1: 19–37. https://doi.org/10.1080/21670811.2014.927986

Li, A., and L. Sinnamon. 2024. "Generative AI Search Engines as Arbiters of Public Knowledge: An Audit of Bias and Authority." *Proceedings of the Association for Information Science and Technology* 61, no. 1: 205–217.

Liang, P., R. Bommasani, T. Lee, et al. 2022. "Holistic Evaluation of Language Models." *arXiv preprint arXiv:2211.09110*. https://doi.org/10.48550/arXiv.2211.09110.

Lin, C. 2022, December 5. "How to Easily Trick OpenAI's genius new ChatGPT." https://www.fastcompany.com/90820252/how-to-easily-trick-openais-genius-new-chatgpt.

Margolin, D. B. 2019. "Computational contributions: A Symbiotic Approach to Integrating Big, Observational Data Studies into the Communication Field." *Communication Methods and Measures* 13, no. 4: 229–47.

Mayson, S. G. 2018. "Bias in, Bias Out." *Yale Law Journal* 128: 2218.

Messeri, L., and M. J. Crockett. 2024. "Artificial Intelligence and Illusions of Understanding in Scientific Research." *Nature* 627: 49–58. https://doi.org/10.1038/s41586-024-07146-0.

Minnesota Journalism Center. 2025, April 24. "Report: As Newsrooms Look to Innovate with AI, Americans Remain Skeptical and Distrusting." https://hsjmc.umn.edu/news/2025-04-24-report-newsrooms-look-innovate-ai-americans-remain-skeptical-and-distrusting.

Mitchell, M. 2019. *Artificial Intelligence: A Guide for Thinking Humans*. Pelican.

Moser, C., T. Lüders, and M. Leidecker-Sandmann. 2024. "AI Discourse in the News Media." *Annual Conference of the Science Communication Division of the German Communication Association 2024*, Parallel Panel I.

Møller, L. A., A. van Dalen, and M. Skovsgaard. 2024. "A Little of that Human Touch: How Regular Journalists Redefine Their Expertise in the Face of Artificial Intelligence. " *Journalism Studies* 26: 84–100. https://doi.org/10.1080/1461670X.2024.2412212.

Napoli, P. M. 2014. "Automated Media: An Institutional Theory Perspective on Algorithmic Media Production and Consumption." *Communication Theory* 24, no. 3: 340–60. https://doi.org/10.1111/comt.12039.

Narayanan, A., and S. Kapoor. 2023, August 18. Does ChatGPT Have a Liberal Bias?. *AI as Normal Technology*. https://www.normaltech.ai/p/does-chatgpt-have-a-liberal-bias

Nishal, S., and N. Diakopoulos. 2024. "Envisioning the Applications and Implications of Generative AI for News Media." *arXiv preprint arXiv:2402.18835*. https://doi.org/10.48550/arXiv.2402.18835.

Ouyang, L., J. Wu, X. Jiang, et al. 2022. "Training Language Models to Follow Instructions with Human Feedback." *Advances in Neural Information Processing Systems* 35: 27730–44.

Paik, S. 2023. "Journalism Ethics for the Algorithmic Era." *Digital Journalism*. https://doi.org/10.1080/21670811.2023.2200195.

Peng, T. Q., and X. Yang. 2025. "Recalibrating the Compass: Integrating Large Language Models into Classical Research Methods." *arXiv preprint arXiv:2505.19402*. https://doi.org/10.48550/arXiv.2505.19402.

Perez, E., S. Ringer, K. Lukošiūtė, et al. 2022. "Discovering Language Model Behaviors with Model-Written Evaluations." *arXiv preprint arXiv:2212.09251*. https://doi.org/10.48550/arXiv.2212.09251.

Peyré, G., and M. Cuturi. 2019. "Computational Optimal Transport: With Applications to Data Science." *Foundations and Trends® in Machine Learning* 11, no. 5–6: 355–607.

Piasecki, S., S. Morosoli, N. Helberger, and L. Naudts. 2024. "AI-generated Journalism: Do the Transparency Provisions in the AI Act Give News Readers What They Hope For?." *Internet Policy Review* 13, no. 4: 1–28. https://doi.org/10.14763/2024.4.1810.

Porlezza, C., and G. Ferri. 2022. "The Missing Piece: Ethics and the Ontological Boundaries of Automated Journalism." *Official Research Journal of the International Symposium on Online Journalism* 12, no. 1: 71–98.

Porto, M. P. 2007. "Frame Diversity and Citizen Competence: Towards a Critical Approach to News Quality." *Critical Studies in Media Communication* 24, no. 4: 303–21. https://doi.org/10.1080/07393180701560864.

Przeworski, A., and H. Teune. 1970. *The Logic of Comparative Inquiry*. Wiley.

Reed, G. F., F. Lynn, and B. D. Meade. 2002. "Use of Coefficient of Variation in Assessing Variability of Quantitative Assays." *Clinical and Vaccine Immunology* 9, no. 6: 1235–39. https://doi.org/10.1128/CDLI.9.6.1235-1239.2002.

Ronanki, K., B. Cabrero-Daniel, J. Horkoff, et al. 2024. "Requirements Engineering Using Generative AI: Prompts and Prompting Patterns." In *Generative AI for Effective Software Development* (pp. 109–127), edited by A. Nguyen-Duc. Springer.

Santurkar, S., E. Durmus, F. Ladhak, C. Lee, P. Liang, and T. Hashimoto. 2023, July. "Whose Opinions do Language Models Reflect?." In *Proceedings of Machine Learning Research*, (pp. 29971–30004) edited by A. Krause, E. Brunskill, K. Cho, B,. Engelhardt, S. Sabato, and J. Scarlett. PMLR.

Shoemaker, P. J., and S. D. Reese. 2014. *Mediating the Message in the 21st Century: A Media Sociology Perspective* (3rd ed.). Routledge.

Shoemaker, P. J., and T. P. Vos. 2009. *Gatekeeping Theory*. Routledge.

Silver, E. 2025, May 7. "Keeping Track of AI use Cases in the Newsroom: A Practical Guide to How Journalists are Using Generative AI and How to Keep Up." https://generative-ai-newsroom.com/keeping-track-of-ai-use-cases-in-the-newsroom-1811b8cb606f.

Simon, F. M., and L. F. Isaza-Ibarra. 2023. "AI in the News: Reshaping the Information Ecosystem?" *Oxford Internet Institute*.

Simon, F. M. 2024. "Escape Me if You Can: How AI Reshapes News Organisations' Dependency on Platform Companies." *Digital Journalism* 12, no. 2: 149–70. https://doi.org/10.1080/21670811.2023.2287464.

Spencer, C. 2025, March 20. "Why Story Discovery is the Killer App for Generative AI in Journalism." https://generative-ai-newsroom.com/why-story-discovery-is-the-killer-app-for-generative-ai-in-journalism-ff49ba2ddef0.

Thäsler-Kordonouri, S., and M. Koliska. 2025. "Journalistic Agency and Power in the Era of Artificial Intelligence." *Journalism Practice* 19, no. 10: 2189–2208. https://doi.org/10.1080/17512786.2025.2480238

Thomson, T. J., R. J. Thomas, M. Riedlinger, and P. Matich. 2025. *Generative AI and journalism: Content, Journalistic Perceptions, and Audience Experiences*. RMIT University.

Thurman, N., S. C. Lewis, and J. Kunert. 2019. "Algorithms, Automation, and News." *Digital Journalism* 7, no. 8: 980–92. https://doi.org/10.1080/21670811.2019.1685395.

Trhlik, F., and P. Stenetorp. 2024. "Quantifying Generative Media Bias with a Corpus of Real-world and Generated News Articles." *arXiv preprint arXiv:2406.10773*. https://doi.org/10.48550/arXiv.2406.10773.

Tolochko, P., H. G. Boomgaarden, and H. Song. 2025, June. "Toward More Robust and Valid Text Annotations by Large Language Models: A Monte Carlo Simulation-based Assessment on Their Validity and Bias." *The 75rd Annual Conference of the ICA. Computational Methods Division*, Denver, CO.

Törnberg, P. 2023. "How to use LLMs for Text Analysis." *arXiv preprint arXiv:2307.13106*. https://doi.org/10.48550/arXiv.2307.13106.

Ulken, E. 2022, December 16. "Generative AI Brings Wrongness at Scale." https://www.niemanlab.org/2022/12/generative-ai-brings-wrongness-at-scale/.

Van Dalen, A. 2024. "Revisiting the Algorithms Behind the Headlines: How Journalists Respond to Professional Competition of Generative AI." *Journalism Practice* 20, no. 3: 897–914. https://doi.org/10.1080/17512786.2024.2389209.

van Rooyen, M. 2013. "Structure and Agency in News Translation: An Application of Anthony Giddens' Structuration Theory." *Southern African Linguistics and Applied Language Studies* 31, no. 4: 495–506. https://doi.org/10.2989/16073614.2013.864445.

Vijay, S., A. Priyanshu, and A. R. KhudaBukhsh. 2024. "When Neutral Summaries Are Not That Neutral: Quantifying Political Neutrality in LLM-Generated News Summaries." *arXiv Preprint, arXiv:2410.09978*. https://arxiv.org/abs/2410.09978.

Villani, C. 2009. *Optimal Transport: Old and New* (Vol. 338, p. 23). Springer.

Wahl-Jorgensen, K., and T. Hanitzsch (eds.). 2019. "Journalism Studies: Developments, Challenges, and Future Directions." In *The Handbook of Journalism Studies* (pp. 3–20). Routledge.

Wolfgang, J. D., T. P. Vos, K. Kelling, and S. Shin. 2021. "Political Journalism and Democracy: How Journalists Reflect Political Viewpoint Diversity in Their Reporting." *Journalism Studies* 22, no. 10: 1339–57. https://doi.org/10.1080/1461670X.2021.1952473.

Wu, S. 2024. "Journalists as Individual Users of Artificial Intelligence: Examining Journalists' 'Value-Motivated Use' of ChatGPT and Other AI tools Within and Without

the Newsroom." *Journalism*. Published ahead of print November 20. https://doi.org/10.1177/14648849241303047.

Zhao, C., Z. Tan, C. W. Wong, X. Zhao, T. Chen, and H. Liu. 2025. "Scale: Towards Collaborative Content Analysis in Social Science With Large Language Model Agents and Human Intervention." *arXiv preprint arXiv:2502.10937*. https://doi.org/10.48550/arXiv.2502.10937.

Ziems, C., W. Held, O. Shaikh, et al. 2024. "Can Large Language Models Transform Computational Social Science?." *Computational Linguistics* 50, no. 1: 237–91.

Author Biographies

Taewoo Kang is an incoming Assistant Professor of Communication in the Department of Communication at Yonsei University, South Korea. His research focuses on digital platformization and political communication. He is currently examining the potential of AI systems to promote deliberative democracy.

Tim Vos is Professor and Director of the School of Journalism at Michigan State University, USA. He is an ICA Fellow and a past president of AEJMC. His research examines the roles of journalism, media sociology and gatekeeping, and media history.

Thomas Hanitzsch is Professor of Communication in the Department of Communication Studies and Media Research at LMU Munich. A former journalist, his teaching and research focus on global journalism cultures, the transformation of journalism, and comparative methodology. He is a Fellow and currently President of the International Communication Association (ICA).

Neil Thurman is Professor in the Department of Media and Communication at LMU Munich and Honorary Senior Research Fellow in the Department of Journalism at City St George's, University of London. His research focuses on the impact of digitization, automation, AI, and the platform economy on media work, content, and audiences. He is author of Media Change: Contemporary Cases, Consequences, and Conceptualizations (Wiley, 2026).

Imke Henkel is a Lecturer (Assistant Professor) in Journalism and Media at the University of Leeds. Imke's research interests include the intersection of journalism and democracy, disinformation, and journalists' professional roles. She is a Worlds of Journalism Study Co-Investigator for the UK and a member of the UK team of the Journalistic Role Performance Project.

Sina Thäsler-Kordonouri is a doctoral candidate and researcher in the Department of Media and Communication (IfKW) at LMU Munich. She is also a Research Fellow at the Public Tech Media Lab from the University of Wisconsin-Madison. Previously, she studied media and communication studies at Freie Universität Berlin and Stockholm University. Her research focuses on the use of AI and automation in journalism as well as related forms of computational journalism.

Wiebke Loosen is a senior researcher at the Leibniz Institute for Media Research | Hans-Bredow-Institut (HBI) (Germany) as well as a professor at the University of Hamburg. Her research focuses on the transformation of journalism within a changing media environment, the journalism-audience relationship, and the automation of societal communication through communicative AI.