From Gatekeepers to Algorithms: Generative AI, News Credibility and Young Adults' Civic Engagement

A conceptual review and research agenda

Debangana Chakrabarti[1]

[1] Assistant Professor, School of Media, Bhawanipur Global Campus, Kolkata, India

ORCID: 0009-0009-2745-214X

Abstract

Generative artificial intelligence (GenAI) is becoming a new intermediary between news and its audiences. Unlike a search engine that points users to journalistic sources, a GenAI system writes a fluent summary and answers the question directly. This paper reviews current scholarship and audience survey evidence on what this shift means for young adults aged roughly 18 to 25, who are the heaviest early adopters. Drawing on gatekeeping theory, mediatization theory and the concept of algorithmic governmentality, it develops a conceptual framework linking four stages: AI-mediated exposure, credibility judgement, verification behaviour and civic engagement. The review identifies a "verification gap", in which the fluency of a generated answer is taken as a sign of its reliability, and argues that this gap may weaken independent checking of claims and reduce exposure to a plural range of journalistic sources. The paper sets out five testable propositions and a research agenda that calls for longitudinal, cross-national and experimental work, since most existing evidence is cross-sectional and drawn from Western countries. It closes with implications for newsrooms, media literacy educators and policymakers.

Keywords: generative AI; journalism; gatekeeping; news credibility; algorithmic trust; young adults; civic engagement; media literacy

1. Introduction

For most of the twentieth century, journalists and editors decided which events became news. Gatekeeping research described this selection process in detail (White, 1950; Shoemaker & Vos, 2009). Online platforms later moved part of that power to ranking and recommendation systems (van Dijck et al., 2018). Generative AI is a third step. A chatbot or AI search summary does not only rank links. It selects, condenses and frames information in a single conversational answer, which performs several tasks that were once the work of reporters and editors (Diakopoulos, 2019).

Audience data show that the shift has begun. In a six-country survey, the share of people who had ever used generative AI rose from 40% in 2024 to 61% in 2025, and seeking information had become the most common weekly use (Fletcher et al., 2025). The Digital News Report 2025 found that 7% of respondents across 48 markets used AI chatbots for news each week, rising to 15% among those under 25 (Newman, 2025). These shares are still small, but they are growing fastest in the age group that will shape future news habits and democratic participation.

The same evidence shows a tension. Audiences use AI for information yet remain wary of AI-made journalism. In the 2025 six-country survey, 62% were comfortable with news produced entirely by a human journalist, compared with 12% for news produced entirely by AI (Fletcher et al., 2025). In the United States, only about a third of those who saw AI-generated search answers said they always or often clicked through to the sources behind them. People distrust AI news in principle but rely on AI answers in practice.

This paper asks how this combination affects young adults' relationship with journalism and public life. It addresses three questions:

1. How does AI-mediated news exposure change young adults' trust in, and judgement of, journalism?
2. How does it affect verification behaviour and exposure to diverse sources?
3. What follows for civic engagement and media literacy?

The paper contributes in three ways. It treats GenAI as a media actor within the news ecosystem, rather than only as an educational or productivity tool. It brings together research from journalism studies, psychology and human-computer interaction that is rarely read together. And it offers a framework and propositions that later empirical studies can test.

2. Review approach

This is a narrative conceptual review, not a systematic one, and it does not report a database search protocol or screening counts. Sources were selected purposively to cover three areas: journalism and gatekeeping theory; research on credibility, automation bias and cognitive offloading; and recent audience surveys on AI and news. Peer-reviewed articles and academic books form the core. Two Reuters Institute reports were added because they hold the most current cross-national data on generative AI and news use, and their figures are used as descriptive evidence only. Each cited source was confirmed to exist through publisher, library or repository records. Because this topic is changing quickly, the review aims to build a framework and agenda, not to count or weigh every study.

3. Theoretical foundations

3.1 Gatekeeping

Gatekeeping theory explains how information is selected and shaped before it reaches an audience (White, 1950). Later work widened the idea to include organisational routines, institutional pressures and social forces (Shoemaker & Vos, 2009). GenAI complicates this picture in two ways. First, the selection is made by a statistical model, not by a person who can be named and held to professional standards. Second, the gate is often invisible: the user sees an answer, not the choices that produced it.

3.2 Mediatization

Mediatization theory holds that media are not only channels but an environment whose logic gradually shapes other social fields (Hjarvard, 2013). Seen this way, AI systems do more than deliver news. They push users toward the formats and rhythms that suit the system, such as short, confident, conversational answers. Over time, expectations of what counts as a good explanation may change to match.

3.3 Algorithmic governmentality

Rouvroy (2013) argues that data-driven systems steer behaviour by anticipating and shaping options, often without prompting conscious reflection. For news, this suggests that the framing and ordering of an AI summary can guide what a young reader notices and what remains unseen, without any explicit persuasion.

3.4 Epistemic trust and cognitive offloading

Two further concepts help explain user behaviour. Epistemic trust is the willingness to rely on others for knowledge. Sahebi and Formosa (2025) argue that AI-mediated communication creates a dilemma: lowering trust protects people from being misled but risks unfair distrust of reliable sources, and it is hard to know which situation applies. Cognitive offloading is the use of external tools to reduce mental effort (Risko & Gilbert, 2016). Offloading is often sensible, but it becomes a concern when the task being offloaded is the evaluation of truth. Related research on automation bias shows that people tend to accept automated recommendations too readily and to check them less often (Parasuraman & Manzey, 2010).

4. What the evidence suggests

4.1 Exposure: from sources to answers

Young adults already rely heavily on social and video platforms for news. In the United States, 54% of 18 to 24-year-olds named these platforms as their main news source in 2025 (Newman, 2025). GenAI adds another layer: the answer may draw on journalism, but the journalist is rarely visible. When users do not click through to sources, the link between a news organisation and its audience weakens, and so does the organisation's ability to earn trust and revenue.

4.2 Credibility: fluency as a cue

Meta-analytic evidence shows that credibility judgements depend on source, message and receiver factors, and that people often use quick cues rather than careful checking (Ou & Ho, 2024). A fluent, well-organised and confident answer is a strong cue of this kind. This review argues that GenAI output supplies such cues at scale, which is why it can be believed even when it is wrong. This is a reasoned inference from the credibility literature and from research on automation bias. It has not yet been tested directly for news summaries, and that gap is the basis of Proposition 1 below.

Audience surveys reveal an apparent paradox. People say they are uncomfortable with AI-made news and expect AI to make news less trustworthy, yet many accept AI answers in daily use (Fletcher et al., 2025). Trust, in other words, is conditional and context-dependent, not stable.

4.3 Verification: the gap between doubt and checking

Survey research on knowledge workers found that higher confidence in GenAI was associated with less self-reported critical thinking, while higher self-confidence was associated with more (Lee et al., 2025). A survey-based study of adults reported a negative association between frequent AI tool use and critical thinking, mediated by cognitive offloading, with younger participants showing higher dependence (Gerlich, 2025). Both studies are cross-sectional and self-reported, so they show association, not cause. They are, however, consistent with the concern that reliance on AI can reduce the effort people put into checking. For this paper the term verification gap describes the distance between the doubt people express about AI and the checking they actually do.

Individual-level interventions can help. A review of misinformation interventions identified several that build resistance, such as lateral reading, prebunking and accuracy prompts, though their effects vary by context and tend to fade without reinforcement (Kozyreva et al., 2024).

4.4 Pluralism and civic engagement

Democratic participation depends on people meeting a range of viewpoints and being able to judge competing claims. If AI summaries compress reporting into one consensus-style answer, minority perspectives and local reporting may drop out, particularly when models are trained on uneven data. At the same time, news avoidance is already high: 40% of respondents across markets said they sometimes or often avoid the news (Newman, 2025). AI summaries could lower the barrier for people who find news overwhelming, which would be a civic benefit. But the same convenience could replace reading journalism with consuming digests. Whether the result is more or less engagement is an open empirical question.

5. A conceptual framework

The framework links four stages. Each stage is shaped by moderators on the user side and the system side.

Stage	Core process	Possible outcome
1. AI-mediated exposure	News encountered through chatbots, AI search summaries and platform assistants	Fewer visits to original journalism
2. Credibility judgement	Fluency, tone and confidence used as quality cues	Trust transferred from institutions to systems
3. Verification behaviour	Checking sources, cross-reading, correcting errors	Gap between stated doubt and actual checking
4. Civic engagement	Political knowledge, discussion, participation	Greater access, or narrower and shallower knowledge

Moderators: need for cognition, prior media literacy, interest in news, education, access to quality journalism, language and cultural context (user side); and source transparency, citation design, error rates and personalisation (system side).

Propositions

P1. The more fluent and confident an AI news answer is, the more credible young adults will judge it, independent of its accuracy.

P2. Frequent reliance on AI answers for news is associated with fewer clicks to original sources and fewer independent verification acts over time.

P3. The link in P2 is weaker among users with high need for cognition and strong media literacy.

P4. Visible source attribution and uncertainty cues reduce the credibility advantage of fluency (P1).

P5. AI-mediated news use predicts civic engagement differently by type: it may raise information-seeking and lower in-depth knowledge and offline participation.

6. Research agenda

Longitudinal designs. Current evidence is mostly cross-sectional. Panel studies that follow the same young adults over several years could show whether dependence on AI grows or whether users learn to calibrate their trust.

Experiments on fluency and error. Controlled tasks in which an AI summary contains a plausible mistake, with measures of detection rate, response time and source visits, would test P1 and P4 directly.

Content analysis. Systematic comparison of AI-generated news summaries with the journalism they draw on, covering framing, attribution, accuracy and range of viewpoints, would show what is lost or changed in summarising.

Beyond the West. The cited audience surveys cover few countries, and most scholarship comes from North America and Europe. Studies in South Asia, Africa and Latin America are needed, where multilingual news, platform use and trust in media differ widely.

Qualitative work. Interviews with young news users, journalists and editors can explain how people understand their own autonomy when answers are automated.

Research ethics. Studies that use deliberately flawed AI output need clear debriefing, data protection and approval by an ethics committee.

7. Implications

For newsrooms. Make attribution and links prominent in any AI-assisted product, state clearly where AI is used, and keep human editorial accountability. Audiences are far more comfortable with human-led news that uses AI assistance than with fully automated news (Fletcher et al., 2025).

For educators. Teach verification as a habit rather than a one-off lesson. Lateral reading and checking a claim against original reporting should be practised with AI tools, not only against traditional misinformation.

For policymakers and platforms. Support transparency about sources, accuracy standards for AI summaries of news, and human oversight for high-stakes information. Policy should aim to strengthen, not substitute for, citizens' capacity to judge information themselves.

8. Limitations

This is a narrative review with purposive selection, so it may omit relevant studies and should not be read as a systematic synthesis. The evidence base on GenAI and news is new and changing quickly, and several studies cited are cross-sectional or self-reported. Audience figures come from a small number of countries and may not describe other settings. The framework is conceptual and the propositions are untested.

9. Conclusion

Generative AI turns news access into a conversation with a system that has no byline and no visible editorial process. For young adults, who adopt it fastest, the risk is not that they will believe everything, but that convenience and fluency may reduce the habit of checking and the contact with varied journalism that informed citizenship needs. The opportunity is real too: easier access to explanations could help people who now avoid the news. Which outcome prevails will depend on how systems are designed, how newsrooms respond, and how well young people are equipped to question what they read. The framework and propositions above offer a starting point for studies that can answer these questions with evidence.

Declarations

Funding. No funding was required.

Conflict of interest. None.

Ethics. This review uses published sources only and involved no human participants.

References

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press. https://doi.org/10.4159/9780674239302

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), Article 6. https://doi.org/10.3390/soc15010006

Hjarvard, S. (2013). *The mediatization of culture and society*. Routledge. https://doi.org/10.4324/9780203155363

Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., Ali, A., Bak-Coleman, J., Barzilai, S., Basol, M., Berinsky, A. J., Betsch, C., Cook, J., Fazio, L. K., Geers, M., Guess, A. M., Huang, H., Larreguy, H., Maertens, R., . . . Wineburg, S. (2024). Toolbox of individual-level interventions against online misinformation. *Nature Human Behaviour, 8*(6), 1044–1052. https://doi.org/10.1038/s41562-024-01881-0

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–22). ACM. https://doi.org/10.1145/3706598.3713778

Newman, N., Ross Arguedas, A., Robertson, C. T., Nielsen, R. K., & Fletcher, R. (2025). *Digital news report 2025*. Reuters Institute for the Study of Journalism. https://doi.org/10.60625/risj-8qqf-jt36

Ou, M., & Ho, S. S. (2024). Factors associated with information credibility perceptions: A meta-analysis. *Journalism & Mass Communication Quarterly, 101*(2), 346–372. [DOI to add from the publisher page]

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Rouvroy, A. (2013). The end(s) of critique: Data-behaviourism versus due-process. In M. Hildebrandt & K. de Vries (Eds.), *Privacy, due process and the computational turn: The philosophy of law meets the philosophy of technology* (pp. 143–167). Routledge. https://doi.org/10.4324/9780203427644

Sahebi, S., & Formosa, P. (2025). The AI-mediated communication dilemma: Epistemic trust, social media, and the challenge of generative artificial intelligence. *Synthese, 205*(3), Article 128. https://doi.org/10.1007/s11229-025-04963-2

Shoemaker, P. J., & Vos, T. P. (2009). *Gatekeeping theory*. Routledge. https://doi.org/10.4324/9780203931653

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

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. https://doi.org/10.1093/oso/9780190889760.001.0001

White, D. M. (1950). The "gate keeper": A case study in the selection of news. *Journalism Quarterly, 27*(4), 383–390. https://doi.org/10.1177/107769905002700403