Journal of Communication and Management

ISSN: 2583-617X (Online)

Journal home page: https://jcoma.com

Research Article

DOI: 10.58966/JCM2025446

Communicating AI Ethics: A Thematic Analysis of Expert Discourse (2020-2025)

Subhodeep Mukhopadhyay*

PhD Scholar, School of Education, GlobalNxt University, Malaysia

ARTICLE INFO

Article history:

Received: 06 November, 2025

Revised: 18 November, 2025

Accepted: 28 November, 2025

Published: 22 December, 2025

Keywords:

Artificial Intelligence, TED Talks, Public Discourse, Science and Technology Communication, AI Ethics, AI Literacy

ABSTRACT

This study explores how expert public discourse addresses key ethical concerns related to artificial intelligence (AI). While much of the existing literature on AI ethics draws on scholarly or policy-based sources, less attention has been paid to how these concerns are communicated to the broader public. TED Talks serve as an important discursive space where complex technological debates are translated for mass audiences by domain experts. Twenty TED Talks (N = 20) on AI literacy and ethics delivered between 2020 and 2025 are examined using inductive thematic analysis. The study identifies recurring issues which are grouped into three core thematic categories: transparency and interpretability, autonomy and power, and structural inequities in AI infrastructure. The first category includes concerns about the opacity of AI systems and the authority granted to machine-generated outputs. The second addresses challenges related to surveillance, human agency, market concentration, and civic participation. The third highlights systemic disparities such as the digital divide, environmental burdens, and control over foundational AI infrastructure. These themes are subsequently interpreted through the lenses of hermeneutics, critical theory of technology, and data justice. The findings contribute to understanding how ethical concerns are framed and narrated in public discourse. They can inform AI literacy initiatives as well as regulatory communication aimed at educators, media professional, and civil society at large.

Introduction

The widespread adoption of artificial intelligence (AI) has brought to fore a number of ethical concerns. Many consider AI as an instrument of algorithmic power. It concentrates authority in the hands of a few large corporations and governments that control the data and resources needed to build advanced models (Crawford, 2021; Zuboff, 2019). This is of course not a new issue. Debates on what *can* and what *should* be delegated to algorithms are quite old. Experts have warned of the dangers of allowing algorithms to mediate social interactions and decision-making (Mittelstadt et al., 2016). Another issue is the opaque nature of algorithms (Pasquale, 2016).This so-called "black-box" problem is especially applicable to AI systems which are often so complex that even their creators cannot explain *how* an AI arrives at a specific decision or response. The quest for equity in AI is another area that has seen a lot of interest. How to stop AI systems from spreading biases continues to be a central theme in the development and governance of AI technologies (Zhou et al., 2024).

Examining the perception of the general public on such ethical concerns is an underexplored area of study, which this study tackles. It examines how experts communicate the ethical aspects of AI through TED Talks, a widely consumed and public-facing medium intended for general audiences. An inductive thematic analysis was done on twenty (N = 20) TED Talks between 2020 and 2025. There are two research questions this study tries to answer:

(1) What are some recurring themes that can be identified in expert discourse around the ethical challenges of AI?

***Corresponding Author:** Subhodeep Mukhopadhyay

Address: PhD Scholar, School of Education, GlobalNxt University, Malaysia

Email ✉**:** sm19184@campus.globalnxt.edu.my

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

(2) How can these themes be interpreted in light of relevant theoretical perspectives?

Literature Review

AI literacy is often defined as a mix of technical knowledge and ethical awareness (Ng et al., 2021). Existing scholarship usually focuses on how educational programs and professional training modules help individuals become more competent in using AI (Lintner, 2024). Another area of research is the integration of AI in media and communication (Chan-Olmsted, 2019). While the ethics of AI has attracted growing scholarly attention (Trattner et al., 2022), there is still limited exploration of how these ethical concerns are represented in public discourse or mass communication—especially from the perspective of AI literacy (Mukhopadhyay, 2025; Ouchchy et al., 2020). In many cases, AI literacy programs do not touch upon these social factors. More often than not, ethics is not covered directly (Kong et al., 2023).

Both AI literacy and AI ethics are evolving fields which have received significant attention in recent years (Batool et al., 2025; Sparks et al., 2024). However, there is limited research on how ethical concerns are communicated in popular formats meant for non-technical people. This is where TED Talks play an important role. Over the years they have emerged as a powerful medium of mass communication, covering a large number of topics, shaping mainstream discourse (MacKrill et al., 2021). TED as a platform provides a space where experts can communicate complex ideas in a way that is easily understandable. In the context of AI, it is like a new arena of AI literacy intervention. Not only can it shape how people make sense of AI, but also helps communicate its ethical and societal implications in an accessible way.

Methodology

Given the exploratory nature of the topic, inductive thematic analysis was used. The aim was to generate data-driven insights to examine how domain specialists communicate different dimensions of AI ethics to non-technical people. TED Talks were selected as the primary data source. Informal YouTube creator channels or academic lectures may not always be able to bridge the gap between experts and the general public, and present complex ideas in simple ways without making them too simplistic. Unlike TED talks, YouTube videos may not always have a controlled and consistent setting, thereby making it difficult to isolate expert-driven narratives. Academic lectures are, of course, more rigorous. However, more often than not they are discipline-specific and less oriented towards the wider public. This makes them unsuitable for examining public-facing framing.

The inclusion criteria for the videos were as follows: (i) the video had to be a TED or TED-affiliated talk between 2020 and 2025 on YouTube; (ii) the central theme must relate to general AI literacy and ethics; (iii) the talk must be in a speech or presentation format; and (iv) the video must have received a minimum of 1,000 views to ensure a baseline level of public engagement. Those videos which did not meet these criteria were removed. This meant that videos on domain-specific applications like healthcare, education, and business solutions, or formats like Q&A sessions, interviews, and panel discussions, were left out. The timeframe 2020 to 2025 was chosen on purpose because during this period, humanity witnessed rapid advancements in AI capabilities, including launch of revolutionary generative AI products.

Initially, 26 videos were shortlisted based on these ideas. Each video was transcribed from YouTube, error-corrected and subsequently analyzed using inductive thematic analysis (Braun & Clarke, 2006). Manual coding was done for the analysis. Although no software like NVivo or ATLAS.ti was used, the study followed a consistent coding protocol. During the initial coding the transcripts were read thoroughly and recurring ideas and metaphors about AI were labeled. These codes were then refined and grouped into code-clusters. The code-clusters or subthemes were then grouped into broader themes. The entire process was iterative. Once a single video was completed, the next video was taken up.

Thematic saturation was observed at the 16th video. At this point no new codes or conceptual ideas were identifiable. To ensure rigor, four more videos were also examined. These videos confirmed the initial findings. No new themes or subthemes were identified, and coding was then concluded. The final themes were formalized for interpretation through the selected theoretical lenses of hermeneutics, critical theory of technology, and data justice. A total of 20 videos (N = 20) was therefore the final data set for analysis.

Results

The analysis identified three major themes that characterize how experts communicate AI ethics: transparency and interpretability, autonomy and power, and structural inequities in AI infrastructure (Table 1). These are discussed in later sections.

Transparency and Interpretability

Transparency and interpretability was identified as a key theme. It covers issues like algorithmic bias, the opacity of AI decision-making, and misplaced trust in machine outputs. These issues reveal how AI systems obscure their inner workings while projecting objectivity. A subset of codes and illustrative quotes are presented in Table 2. The subthemes are discussed subsequently.

Embedded Biases

A recurring theme across many of the talks was the presence of deep-seated biases within AI systems which

Table : 1 Key Themes

Theme	*Content Description*
Transparency and Interpretability	Concerns about the opacity of AI systems and the undue authority granted to machine-generated outputs.
Autonomy and Power	Challenges related to surveillance, erosion of human agency, market concentration, and threats to civic participation.
Structural Inequities in AI Infrastructure	Systemic disparities including the digital divide, environmental burdens, and control over foundational AI infrastructure.

Table : 2 Transparency and Interpretability: Subthemes with Representative Codes and Quotes

Subtheme	*Code*	*Quote*	*Source*
Embedded Biases	Subconscious bias in AI	"Bias can be embedded in AI on a subconscious level"	(TEDx Talks, 2024a)
	Biased training data	"If the AI is only trained on ... raw web data, which is freely available, it's not good because this data is loaded with racism and sexism and misinformation."	(TED, 2023a)
	Masculine authority bias	"…male AI assistants are considered better suited for roles that require authority and leadership"	(TEDx Talks, 2024a)
	Racial bias in facial recognition	"...common facial recognition systems were vastly worse for women of color compared to white men"	(TED, 2023b)
The "Black Box" Problem	Complexity of model architecture	"The sheer size of these neural networks makes it difficult for us to see how they are… making decisions"	(TEDx Talks, 2024c)
	Limits of explainability	"What makes it hard for experts to know what's going on is basically just, there are too many numbers…"	(TED, 2024b)
Epistemic authority of AI	Hallucination framed as fact	"AI could confidently state fiction as fact"	(TEDx Talks, 2024a)
	Human origin of AI outputs	"…it was a human that collected the data, and it was a human that trained and tested the results"	(TEDx Talks, 2023)
	Prompting user skepticism	"Knowing when to abstain from using these tools is just as vital, and prepares you to make informed decisions"	(TEDx Talks, 2024e)

manifest in various ways including racism, sexism, and misinformation. The speakers provide several examples. The classic example is that of AI voice assistants where female personas are associated with support roles, while male voices are associated with authority and leadership. Another example that was discussed was facial recognition. In the Western world, these systems have been shown to affect women of color more than white men. A further challenge was policymakers limited grasp of AI, especially its ethics. This often becomes a roadblock to drafting meaningful regulations.

The "Black Box" Problem

The internal working of AI models is quite complex. Even experts cannot often interpret or understand how AI systems arrive at a decision. This is commonly called the "black box" problem. This raises important questions about accountability and trust.

Epistemic authority of AI

This theme is based on the idea that many users tend to view AI outputs as trustworthy simply because a machine produced them. Issues like inaccurate output, hallucinations and the human element in the AI value chain are seemingly disregarded. The speakers emphasized the need for users to question AI outputs and understand their boundaries.

Autonomy and Power

This theme is about how AI systems affect individual freedom and centralize power, thereby weakening democratic processes. Findings suggest a constant tension between human agency and corporate control. A subset of codes along with representative quotes and sources is presented in Table 3.

Data Use and User Autonomy

Several speakers drew attention to the pervasive use by AI-driven systems of personal data ranging from browsing patterns to online interactions. While some suggested that this could potentially improve efficiency, the general view was critical of how such data was routinely exploited. Concerns were raised about manipulative design features that leverage behavioral data to increase engagement. In particular, profiling vulnerable populations, such as children, was highlighted. Experts also highlighted

Table : 3 Autonomy and Power: Subthemes with Representative Codes and Quotes

Subtheme	*Code*	*Quote*	*Source*
Data Use and User Autonomy	Data-driven manipulation	"But they also use… the data they gather about you to build more addictive products for you, to target your vulnerabilities"	(TEDx Talks, 2020a)
	Opaque data circulation	"These kinds of uses of AI — the ones you may not see every day — are still impacting your life…"	(TEDx Talks, 2020c)
	Time-saving potential	"Just imagine a time… when every one of us will have so much artificial intelligence around us that it will take care of the routines, the things we're not good at, that we don't like, or that just consume so much time."	(TEDx Talks, 2025a)
Concentration of Power	Resource imbalance	"…extreme-scale AI models are so expensive to train, and only a few tech companies can afford to do so."	(TED, 2023a)
	Power disparity concern	"…who would you trust to have a million times more power and wealth than any other actor in society? Any company? Any government? Any individual?"	(TED, 2025a)
	Sustainable AI engineering	"It has one deceptively simple mission: to establish a new branch of engineering to take AI safely, sustainably and responsibly to scale."	(TED, 2021)
Importance of Democratic Control	AI vs democracy	"…a lot of people think AI is toxic for democracy—because it enables the spread of misinformation and disinformation; it ferments political polarization."	(TEDx Talks, 2025b)
	Safety vs market dominance	"…the more shortcuts you take to get market dominance or prove you have the latest capabilities, the more money you can raise, the more ahead you are in the race"	(TED, 2025a)
	Lack of regulation	"a sandwich has more regulation than AI"	(TED, 2025b)
Surveillance	Data sold to third part	"An AI-literate citizen is someone who… recognizes how their personal data is being stored, shared, deleted, and sold to a third party"	(TEDx Talks, 2023)
	Lethal consequences	"…wrong to build technologies for military surveillance with potentially lethal outcomes"	(TEDx Talks, 2020a)
Risks of AI with greater Agency	Deploying unsafe AI	"we're currently releasing the most powerful, inscrutable, uncontrollable technology we've ever invented that's already demonstrating behaviors of self-preservation and deception that we only saw in science fiction movies"	(TED, 2025a)

the opaque nature of corporate data flows. At the same time, a few speakers noted that AI systems could relieve individuals from mundane or repetitive tasks and help one focus on higher order tasks.

Concentration of Power

A recurring theme was the growing concentration of AI development capacity in the hands of a few large companies. The costs required to train advanced models are huge. Smaller entities therefore find it difficult to participate in AI research. This gives rise to the metaphor of a "dystopia" - a future where a few companies hoard power and resources.

Importance of Democratic Control

Speakers emphasized growing concerns around AI's influence on democratic processes due to accelerating misinformation and deepening political polarization. Rapid commercialization of AI, driven by incentives for market dominance, was seen as another major risk, since safety and accountability are often sidelined in the quest for growth. One speaker observed how even a mundane consumer good like a sandwich receives more regulatory oversight than AI systems.

Surveillance

Another frequently raised concern was the use of AI in surveillance. User activities are continuously tracked, often without consent. Their data too is shared with third parties. AI is increasingly being used to monitor and control how people behave. Speakers also drew attention to high-risk applications, such as military-grade drone surveillance. These developments were interpreted as symptoms of a broader erosion of civic space.

Risks of AI with greater Agency

Experts highlighted mounting commercial pressure to develop AI systems capable of independent decision-

Table : 4 Structural Inequities in AI Infrastructure: Subthemes with Representative Codes and Quotes

Subtheme	*Codes*	*Example Snippets*	*Source*
AI as Infrastructure and Energy as Currency	AI as critical infrastructure	"…if you look from a top-down perspective, that AI is not just a technology; it's infrastructure."	(TEDx Talks, 2024f)
	Energy-intelligence inequality	"the more energy you have, the more access to intelligence you have"	(TEDx Talks, 2024f)
Digital Divide	Paid vs. free access divide	"…the free AI is not the AI that's going to be the best version for our students. It's going to be the one that now costs money, and if there isn't equitable access, then we create a digital divide…"	(TEDx Talks, 2024b)
	Technical Expertise	"…ability to do a base amount of programming can be quite impactful"	(TEDx Talks, 2022)
Environmental Inequity	Ecological cost of large models	"…just training it used as much energy as 30 homes in a whole year and emitted 25 tons of carbon dioxide…"	(TED, 2023b)
	Larger not always better	"…switching out a smaller, more efficient model for a larger language model emits 14 times more carbon for the same task. Like telling that knock-knock joke"	(TED, 2023b)

making and task execution. This was seen as deeply problematic in the absence of adequate understanding or safeguards. Concerns were voiced over a "race to roll out" scenario, where safety considerations are bypassed in favor of speed and profit. Some speakers warned that we may be unleashing technologies with behavioral traits reminiscent of self-preservation or deception.

Structural Inequities in AI Infrastructure

Speakers emphasized how AI development amplifies existing inequalities. This is because of uneven access to energy and advanced tools, and consequently those with more resources gain disproportionate benefits. Meanwhile marginalized groups often bear the ecological costs. A subset of codes, representative quotes, and sources is presented in Table 4.

AI as Infrastructure and Energy as Currency

AI is being seen as the next great infrastructure fueling the growth of new economies, similar to roads and the internet of the previous era. Energy and data are the currency of this new infrastructure. It was noted with great concern that many companies are already securing direct energy sources, and that such trends are likely to deepen existing global inequities. Greater AI capabilities would therefore accrue mostly to only energy-rich actors.

Digital Divide

Speakers highlighted the growing risk of an AI-induced digital divide, particularly as advanced tools become increasingly pay-walled or have limited capabilities in free versions. Those without premium access could be systematically disadvantaged.

Environmental Inequity

Another important theme highlighted was the environmental impact of AI, a topic which has seen considerable interest in recent times. These include energy use and emissions that are often underreported or not publicly audited. Vulnerable communities are likely to bear a disproportionate share of this burden, despite having little influence over the technologies in question.

Analysis

Expert public-facing discussions on platforms like TED differ greatly from academic or policy discussions in structure and tone. The target audience is quite different. Unlike academic publications, which follows a specific pattern and rely on rigorous methodology, TED Talks are narrative-driven and designed for wide dissemination. They aim to persuade and simplify. They are not intended to exhaustively argue or prove something. The purpose is to explain complex ethical and political concerns in ways that can be easily understood. This makes them a powerful cultural instrument (MacKrill et al., 2021).

Let us take the example of AI opacity. An academic paper may explain the concept formally in terms of explainability and algorithmic accountability (Burrell, 2016). However, a TED speaker explains the very same idea through the metaphor: "I don't understand AI, and neither does anyone else" (TED, 2024b). "Black box" therefore ceases to be a technical term. Rather it assumes the role of a rhetorical device, one that emphasizes uncertainty and loss of control in an AI-dominated world. TED Talk discourse is in essence affective and interpretive. The findings of the previous sections must be viewed against this backdrop.

To understand their implications, the identified themes have to be looked at through relevant theoretical perspectives. Transparency and interpretability can be explained through hermeneutics. Critical theory of technology can well explain autonomy and power. Theory of data justice allows us to understand structural

Table : 5 Themes, Categories and Theoretical Perspectives

Theme	*Categories*	*Theoretical Perspectives*
Transparency and Interpretability	Embedded Biases	Hermeneutics
	Black-box problem	
	Epistemic Authority of AI Output	
Autonomy and Power	Data Use and User Autonomy	Critical Theory of Technology
	Concentration of Power	
	Importance of Democratic Control	
	Surveillance	
	Risks of AI with Greater Agency	
Structural Inequities in AI Infrastructure	AI as Infrastructure and Energy as Currency	Data Justice
	Digital Divide	
	Environmental Inequity	

inequities in AI infrastructure (Table 5). This allows us to relate the findings to broader conceptual traditions. Without these theoretical underpinnings, we will not be in a position to decode how public discourse reflects and reshapes our normative orientation toward AI.

Hermeneutic View of AI Transparency and Interpretability

In the Gadamerian tradition, understanding is never passive but dialogic emerging as it does from the interaction of the topic and prior assumptions, or what Gadamer refers to as *vorurteile* (Schmidt, 2014). Seen from this angle, AI-chat may appear to be a hermeneutic process. The user begins with their own assumptions or understanding, while the AI responds with information from its vast "horizon" of data on which it has been trained. During this exchange, AI systems bring their own underlying assumptions. They also have embedded prejudices in terms of gender, race and worldviews, an issue which has been discussed extensively in literature (Livingston, 2020; Mukhopadhyay & Reddy, 2023; O'Connor & Liu, 2024). This man-machine interaction seemingly creates a hermeneutic circle. The user refines their query based on the AI's answer, and vice versa, and the cycle goes on till the user is satisfied that his queries have been answered.

However, a true fusion of horizons does not happen. This is because AI lacks a subjective experience. It is not sentient, and therefore, the process of "meaning-making," where a mutual understanding between two conscious beings develops, never materializes. The AI may appear to be aware. But this awareness is simply statistical rules. It is not genuine understanding, as has been pointed out by a number of speakers (TEDx Talks, 2024d). What AI produces is not true knowledge as we humans understand the word, and hence should be used mostly to augment one's capabilities and not to bypass the learning process itself (TEDx Talks, 2024b). All this is further complicated by AI's opacity. The latter is not just a technical barrier but a crisis in how knowledge is constructed and shared, a veritable epistemological rupture. The situation grows more worrisome when one considers AI's capacity for recursive self-improvement, enabling them to potentially refine their code completely bypassing humans in the loop (TED, 2024a).

Therefore, when AI experts declare that neither they nor others actually understand how AI works (TED, 2024b), it reflects a breakdown in our ability to comprehend the very systems we create. We simply do not understand how AI arrives at a decision. We have no way of knowing, at least at this point in time, why AI systems provide the answers that they do. Ontologically, AI is not a mere tool or utilitarian object in the world but becomes a quasi-agentic system. It is akin to a presence that participates in shaping social reality without being fully comprehensible. The metaphor of AI as a "country full of geniuses in a data center" (TED, 2025a), exemplifies the scale at which AI output exceeds the threshold of human-cognition. Seen in this light, AI is far from a neutral agent. There is indeed a pressing need to "ask questions about [AI's] safety, trustworthiness, and limitations" (TEDx Talks, 2024e).

Autonomy and Power in AI: Critical Perspective

Concerns on data use, autonomy, power, and surveillance point to an evolving AI ecosystem that disrupts democratic norms. AI is no longer a passive tool. Rather, it is actively shaping and also being shaped by society. User exploitation and monopolistic control by tech giants are real issues in the public consciousness. Lack of regulatory oversight, and the erosion of public trust in AI-enable democratic processes further complicate matters. It is not surprising that ethicists, futurists, policy-makers and public intellectuals are demanding a critical scrutiny of the forces behind global AI systems. With increased usage in critical areas like healthcare, warfare, policing, and justice, questions naturally arise about who wields the power. The critical theory of technology therefore becomes an appropriate lens to understand the implications of increasing autonomy of AI systems and the ostensible concentration of power in the hands of a few. According to

the critical view, technology is not neutral. Neither is it an autonomous force that dictates social change (Feenberg, 2008). AI tools reflect existing hierarchies and end up reinforcing inequalities.

As noted by the speakers, they are typically developed by elite tech actors who prioritize profit over public good. Another concern is AI's role in amplifying misinformation and political polarization. There are concerns about a growing technocracy at the helm of which are unelected experts. Seen in this light, AI is akin to a battleground where power and agency are contested and reshaped. It is the arena where the elites fight the elites for AI market domination, while the gaps between the "haves" and "have-nots" increase (Malhotra, 2021). Many of the speakers have cautioned against this dangerous trend. They have asked for more scrutiny and better governance of AI systems, which may be viewed in the light of what Feenberg refers to as "democratic rationalization", or the opening up the development process to broader public scrutiny and participation.

AI Structural Inequities and Data Justice

Viewed through the lens of data justice (Taylor, 2017), AI systems reveal and reproduce deep structural inequalities. This perspective lends credence to the views of the speakers that the material and political conditions of AI production disproportionately benefit a handful of well-resourced actors. Data must not be treated as abstract artefact. The resource-intensive nature of training large-scale models naturally concentrates AI development within a few powerful tech corporations, mostly in the Global North. These entities control the development of large-scale AI models. More importantly, they also monopolize the underlying infrastructures, the data centers and energy sources, needed to sustain complex AI systems. Independent researchers and the civil society actors have little to no voice in the development of AI. Data justice also foregrounds the importance of inclusion in design processes by asking questions like whose data is being used, who benefits from AI, who bears its environmental costs, and who solves its social problems. Calls for "diverse teams with different backgrounds, cultures, and skills" (TEDx Talks, 2024a) aim to ensure that AI systems better reflect social complexity. The idea of data justice allows us to draw interesting parallels to this trajectory of AI and colonialism, giving rise to the idea of digital colonialism, where Global North AI infrastructures extract data and resources from the Global South without equitable benefit-sharing. Addressing such injustices requires confronting algorithmic inequities and surveillance capitalism (Zuboff, 2019).

Conclusion

The study focused on TED and TED-affiliated talks, and excluded other platforms and formats like academic lectures, podcasts, and general informational videos. Moreover, only formal presentations were analyzed. Q&A sessions, panel discussions, interviews, and other interactive formats were not considered. They could have offered different insights. Another shortcoming is that the study treats the period 2020 to 2025 as one singular block and does not look at how the discourse on AI ethics has evolved over the years. New revolutionary GAI products like GPT and recent agentic products are not separately accounted for. Future research could address these gaps. They can examine multi-platform and multiformat expert discourse. Comparative studies across media genres and studies on how GAIs, other major technological shifts or political events, may have changed the discourse on AI ethics, would deepen our understanding of the AI ethics landscape.

Despite these gaps, the study is important as it demonstrates that not only is there a growing concern regarding the ethical aspects of AI. Using inductive thematic analysis, this study identified three themes: transparency and interpretability, autonomy and power, and structural inequities in AI infrastructure. The findings were interpreted through the lenses of hermeneutics, critical theory of technology, and data justice. They tell a fascinating tale about the lesser-discussed aspects of AI. Terms like accountability and transparency may sound like jargon, but have real-world implications. While AI may appear to understand what we say and communicate with us in a limited sense, it is to be noted that we are in essence dealing with a black box that is resistant to dialogue and reciprocal understanding.

The study has different takeaways for different groups of people. It is a cautionary tale for software developers. Just because one *can* develop an AI system does not mean that they necessarily *should* – a point also noted by one of the speakers (TEDx Talks, 2020b). We must understand that an AI does not "speak" in human language. Nor does it invite interpretation in traditional ways. Hence, value-sensitive design of AI systems is the need of the hour. This study offers several key points for educators. AI literacy programs should teach not only how AI functions, but also its ethical implications. The fact that AI is also a sociopolitical phenomenon must also be conveyed to students. An examination of AI-generated content and its onto-epistemic implications is also needed. Concepts like opacity, transparency, algorithmic bias, and digital colonialism can add value to classroom discussions on the impact of AI. For policymakers, the study suggests that they go beyond technical regulations. Policies must emphasize civic participation. For media practitioners, the analysis illustrates the importance of moving beyond hype and alarmism. They should instead focus on narratives that tell us how AI can be used for social good.

References

1. Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. *AI and Ethics*, *5*(3), 3265–3279. https://doi.org/10.1007/s43681-024-00653-w
2. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.

org/10.1191/1478088706qp063oa
3. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 2053951715622512. https://doi.org/10.1177/2053951715622512
4. Chan-Olmsted, S. M. (2019). A Review of Artificial Intelligence Adoptions in the Media Industry. *International Journal on Media Management*, *21*(3–4), 193–215. https://doi.org/10.1080/14241277.2019.1695619
5. Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
6. Feenberg, A. (2008). Critical Theory of Technology: An Overview. In G. J. Leckie & J. E. Buschman (Eds.), *Information Technology in Librarianship: New Critical Approaches* (1st ed., pp. 31–46). www.lu.com. https://doi.org/10.5040/9798400670268
7. Kong, S.-C., Cheung, W. M.-Y., & Tsang, O. (2023). Evaluating an artificial intelligence literacy programme for empowering and developing concepts, literacy and ethical awareness in senior secondary students. *Education and Information Technologies*, *28*(4), 4703–4724. https://doi.org/10.1007/s10639-022-11408-7
8. Lintner, T. (2024). A systematic review of AI literacy scales. *Npj Science of Learning*, *9*(1), 50. https://doi.org/10.1038/s41539-024-00264-4
9. Livingston, M. (2020). Preventing Racial Bias in Federal AI. *Journal of Science Policy & Governance*, *16*(02). https://doi.org/10.38126/JSPG160205
10. MacKrill, K., Silvester, C., Pennebaker, J. W., & Petrie, K. J. (2021). What makes an idea worth spreading? Language markers of popularity in TED talks by academics and other speakers. *Journal of the Association for Information Science and Technology*, *72*(8), 1028–1038. https://doi.org/10.1002/asi.24471
11. Malhotra, R. (2021). *Artificial intelligence and the future of power: 5 battlegrounds*. Rupa.
12. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 2053951716679679. https://doi.org/10.1177/2053951716679679
13. Mukhopadhyay, S. (2025). AI Literacy among Media Professionals in the Age of GPTs. In M. Jana & S. Banerjee (Eds.), *Artificial intelligence and contemporary media* (pp. 45–66). Notion Press.
14. Mukhopadhyay, S., & Reddy, D. (2023). Artificial Intelligence as an Enabler of Western Universalism. In R. Malhotra, T. N. Sudarshan, & M. Sastry (Eds.), *The Power of Future Machines: Essays on Artificial Intelligence* (pp. 241–263). BluOne Ink LLP.
15. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041
16. O'Connor, S., & Liu, H. (2024). Gender bias perpetuation and mitigation in AI technologies: Challenges and opportunities. *AI & SOCIETY*, *39*(4), 2045–2057. https://doi.org/10.1007/s00146-023-01675-4
17. Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI & SOCIETY*, *35*(4), 927–936. https://doi.org/10.1007/s00146-020-00965-5
18. Pasquale, F. (2016). *The black box society: The secret algorithms that control money and information* (First Harvard University Press paperback edition). Harvard University Press.
19. Schmidt, L. K. (2014). *Understanding hermeneutics*. Routledge, Taylor & Francis Group.
20. Sparks, J. R., Ober, T. M., Tenison, C., Arslan, B., Roll, I., Deane, P., Zapata-Rivera, D., Gooch, R. M., & O'Reilly, T. (2024). *Opportunities and Challenges for Assessing Digital and AI Literacies*. ETS Research Institute. https://www.ets.org/pdfs/rd/ets-digital-literacy-ai-full-report.pdf
21. Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2), 205395171773633. https://doi.org/10.1177/2053951717736335
22. TED. (2021, January 15). *6 big ethical questions about the future of AI | Genevieve Bell* [Video]. https://www.youtube.com/watch?v=UGHzKaAOOcA
23. TED. (2023a, April 28). *Why AI Is Incredibly Smart and Shockingly Stupid | Yejin Choi | TED* [Video]. https://www.youtube.com/watch?v=SvBR0OGT5VI
24. TED. (2023b, November 6). *AI Is Dangerous, but Not for the Reasons You Think | Sasha Luccioni | TED* [Video]. https://www.youtube.com/watch?v=eXdVDhOGqoE
25. TED. (2024a, April 22). *What Is an AI Anyway? | Mustafa Suleyman | TED* [Video]. https://www.youtube.com/watch?v=KKNCiRWd_j0
26. TED. (2024b, May 1). *How to Govern AI — Even If It's Hard to Predict | Helen Toner | TED* [Video]. https://www.youtube.com/watch?v=LUn8IjZKBPg
27. TED. (2025a, May 1). *Why AI Is Our Ultimate Test and Greatest Invitation | Tristan Harris | TED* [Video]. https://www.youtube.com/watch?v=6kPHnl-RsVI
28. TED. (2025b, May 21). *The Catastrophic Risks of AI — and a Safer Path | Yoshua Bengio | TED* [Video]. https://www.youtube.com/watch?v=qe9QSCF-d88
29. TEDx Talks. (2020a, February 5). *How To Solve AI's Ethical Puzzles | Cansu Canca | TEDxCambridgeSalon* [Video]. https://www.youtube.com/watch?v=cplucNW70II
30. TEDx Talks. (2020b, February 19). *AI Literacy, or Why Understanding AI Will Help You Every Day | Jordan Harrod | TEDxBeaconStreet* [Video]. https://www.youtube.com/watch?v=cKl9QsVv7hY
31. TEDx Talks. (2020c, October 22). *AI Literacy: What is It and Why Do We Need It? | Nisha Talagala | TEDxHarkerSchool* [Video]. https://www.youtube.com/watch?v=IBWajXiJ434
32. TEDx Talks. (2022, November 22). *The Importance of AI Literacy in Society | Armen Kherlopian | TEDxAUA* [Video]. https://www.youtube.com/watch?v=J87vGLCBXDY
33. TEDx Talks. (2023, October 26). *The AI Literate Global Citizen | Melissa Drew | TEDxHagerstownWomen* [Video]. https://www.youtube.com/watch?v=vzQJC8nKnxQ
34. TEDx Talks. (2024a, March 4). *AI Literacy In Today's World | Parmida Beigi | TEDxSFU* [Video]. https://www.youtube.com/watch?v=6rBYvqBVmGM
35. TEDx Talks. (2024b, March 22). *AI Literacy: The Key to Responsible Use of AI in Education | Mary Lou Maher | TEDxUNCCharlotte* [Video]. https://www.youtube.com/watch?v=weemM1r-zaY
36. TEDx Talks. (2024c, April 16). *Why AI Can't be Ethical—Yet | Eleanor Manley | TEDxDaltVila* [Video]. https://www.youtube.com/watch?v=9DXm54ZkSiU
37. TEDx Talks. (2024d, June 10). *The importance of AI literacy for all | Mike Kentz | TEDxSavannah* [Video]. https://www.youtube.com/watch?v=DRyWDNjRWaw
38. TEDx Talks. (2024e, August 27). *Ethical considerations for generative AI | Sriram Natarajan | TEDxWoodinville* [Video]. https://www.youtube.com/watch?v=EkF2wvflRns
39. TEDx Talks. (2024f, November 13). *The importance of AI literacy | Dominic Boenisch | TEDxSchwabing* [Video]. https://www.youtube.com/watch?v=Qyj6deRlFqc
40. TEDx Talks. (2025a, January 2). *The future of AI, work, and human potential | Lars Thomsen | TEDxHWZ* [Video]. https://www.youtube.com/watch?v=QOKLW5ITEiI
41. TEDx Talks. (2025b, February 16). *Why AI is a threat—And how to use it for good | John Tasioulas | TEDxAthens* [Video]. https://www.youtube.com/watch?v=8vQGGcPojM4
42. Trattner, C., Jannach, D., Motta, E., Costera Meijer, I., Diakopoulos, N., Elahi, M., Opdahl, A. L., Tessem, B., Borch, N., Fjeld, M., Øvrelid, L., De Smedt, K., & Moe, H. (2022). Responsible media technology and AI: Challenges and research directions. *AI and Ethics*, *2*(4), 585–594. https://doi.org/10.1007/s43681-021-00126-4
43. Zhou, M., Abhishek, V., Derdenger, T., Kim, J., & Srinivasan, K. (2024). *Bias in Generative AI* (No. arXiv:2403.02726). arXiv. https://doi.org/10.48550/arXiv.2403.02726
44. Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile books.

HOW TO CITE THIS ARTICLE: Mukhopadhyay, S. (2025). Communicating AI Ethics: A Thematic Analysis of Expert Discourse (2020-2025). *Journal of Communication and Management*, 4(4), 49-56. DOI: 10.58966/JCM2025446