

Artificial intelligence, Algorithmic Power and the Transformation of Democracy

Received 17 July 2026

Acceptance 13 September 2026

Published October 2026

Okil Sunar

MPhil Leading to PhD Scholar

Mid-West University, Surkhet

sunamokil@gmail.com

Abstract

Artificial intelligence (AI) is reshaping the infrastructure of democratic life, simultaneously extending the reach of surveillance and reorganizing the spaces in which citizens deliberate. This paper offers a theoretical synthesis of two traditions rarely placed in sustained dialogue with one another in the AI-and-democracy literature: Michel Foucault's theory of disciplinary power, biopolitics, and governmentality, and Jürgen Habermas's theory of the public sphere and communicative action. Drawing on a systematic review of peer-reviewed and scholarly sources, from an initial pool of 100 identified works twenty were retained for close synthesis, ten engaging Foucauldian surveillance theory and ten engaging Habermasian theories of political participation and deliberation. Six cross-cutting themes emerged: the transformation of the panopticon into algorithmic governmentality; the biopolitical management of populations through predictive data; the entanglement of surveillance capitalism with the colonization of the lifeworld; the algorithmic restructuring of the public sphere; the promise and limits of AI as a deliberative mediator; and the prospects for democratic resistance and accountability. The synthesis suggests that AI does not simply strengthen or simply threaten democracy; it does both at once, disciplining citizens through opaque data infrastructures while also creating new, contested spaces for rational-critical discourse. The paper concludes with implications for democratic theory, technology governance, and future research.

Keywords: *artificial intelligence, democracy, surveillance, Michel Foucault, Jürgen Habermas, public sphere, deliberation, governmentality*

Background

AI technology has evolved from an incidental technical novelty to a core component of politics in the present age. Predictive policing algorithms, facial recognition technology, automated content moderation, algorithmic news filtering and large language models expressly designed to facilitate deliberation have become the central means by which citizens are surveilled, informed and listened to. Two strands of theory in particular provide especially potent and dramatically contrasting frames for interpreting this transformation. The first, which is rooted in Michel Foucault's work, conceives of power as productive and not simply coercive, operating through the production and distribution of knowledge about individuals

and populations in disciplinary institutions, biopower and governmentality, epitomized in the design of the panopticon (Foucault, 1977). The second, which stems from the writings of Jürgen Habermas, regards democratic legitimacy as being grounded in rational-critical debate that takes place in a public sphere insulated from state and market domination, as well as from communicative action aimed at reaching mutual understanding and not manipulating others (Habermas, 1991, 1996).

In this sense, this paper hopes to bring these two traditions into conversation when it comes to thinking about artificial intelligence and democracy. It hopes to inquire into the question of what insights each tradition can offer regarding the way in which AI technologies both surveil citizens and, simultaneously, shape their capacity to deliberate. There has been an explosion of research in the tradition of Foucault concerning AI, in which he re-interprets the idea of the panopticon in a new world of predictive policing, social credit scores, and algorithmic normalizations (Dai et al., 2025; Sahakyan et al., 2025; Varela, 2024). Research in the tradition of Habermas has similarly expanded to meet a different but related set of questions, including not only the fragmentation of the public sphere by new technologies, but also the more novel suggestion that AI itself might be engineered to support conditions of rational-critical deliberation, as in Google DeepMind's so-called Habermas Machine (Jungherr & Schroeder, 2023; Tessler et al., 2024).

But beyond the mere categorization of AI from an academic standpoint, governments, tech companies, and international institutions are increasingly making decisions on regulation, deployment, and evaluation of the democratic effects of artificial intelligence in the absence of an explicit theoretical framework. For instance, the Artificial Intelligence Act of the European Union classifies prohibited uses, high-risk applications, and low-risk applications primarily in relation to harm done to safety and fundamental rights concepts which cannot perfectly overlap with disciplinary normalization as understood by Foucault or discourse legitimacy as understood by Habermas (Sahakyan et al., 2025). Thus, such a theoretical synthesis, as attempted above, can have both theoretical and practical importance. It can allow one to understand clearly what, exactly, is threatened by the various classes of AI applications, whether that threat is the erosion of individual liberty through disciplinary normalization, the erosion of the capacity of self-government of the community through deterioration of public reason, or, as suggested in this paper, the entanglement of both these aspects. Recognizing the nature of the risk involved will determine the appropriate measures which would prove effective to counteract it, whether legal limitation of data collection practices, structural modification of the platform itself, or cultivation of the civic skills needed to criticize algorithms.

Indeed, many of the studies dedicated to artificial intelligence and democracy consider surveillance and deliberation as two distinct research problems, each studied by distinct

groups of scholars who employ distinct vocabularies (Capodivacca & Giacomini, 2024; Volpe, 2025). Surveillance scholars, influenced by Foucault's ideas, emphasize the disciplinary and biopolitical aspects of data collection, analytics, and control, paying insufficient attention to the fact that the same technologies also change the discursive space in which citizens' opinions form and accountability is enforced (Capodivacca & Giacomini, 2024; Gane, 2012). At the same time, deliberation scholars influenced by Habermas focus their research on the normative preconditions for proper public discourse, considering the role that AI can play in ensuring or challenging these preconditions (Habermas, 2022; Volpe, 2025). In other words, there is a notable gap between the two approaches in understanding how the technologies which enable AI surveillance affect public discourse and deliberation. It is nearly impossible to explain how a particular technology (such as social media) can be an instrument of behavior surveillance and at the same time – the infrastructure for deliberation using either of the two approaches separately. This means that the research problem of the paper is synthetic and theoretical rather than empirical it is about reading Foucault and Habermas together in the context of AI in democracy.

Objectives

This research was objectives to review peer-reviewed theoretical literature that utilizes the ideas of Foucault on discipline, biopower, and governmentality on artificial intelligence. To review peer-reviewed literature on theory that applies Habermas's notions of public sphere, communicative action, and deliberative legitimacy to artificial intelligence. To recognize converging and diverging themes between the two areas of research by considering them complementary instead of exclusive perspectives on each other. To consider the significance of this synthesis in relation to democratic theory, technology governance, and future studies in artificial intelligence and politics.

Methodology

The current study is conducted within a qualitative, theory-driven literature synthesis approach. Unlike primary data collection, theory synthesis involves a process where already existing contributions of scholars to a particular topic, in this case peer-reviewed works that make use of surveillance theory by Michel Foucault and Habermas' theory of the public sphere and communicative action, are collected in order to determine areas where these approaches coincide and differ from each other in order to build an integrated model. In addition to this, the current research does not use a quantitative, statistical approach, but an interpretive one – literature was not summarized numerically, but rather conceptually, based on themes derived inductively from literature. Such an approach is relevant in light of the problem statement of this paper, since it is defined theoretically: the question lies in the interpretation of theories, not people's behavior under certain conditions.

Three types of sources were used. The first type of sources, and the main body of

evidence on which the synthesis was based, included peer-reviewed journal articles using Foucauldian or Habermasian theory in relation to AI, algorithmic governance, surveillance, or digital deliberation. The second type of sources included original works by Foucault and Habermas, which provided a grounding of the secondary literature in an accurate understanding of the theories. The third type of sources included closely related bridging works, such as the theory of surveillance capitalism formulated by Zuboff (2019), that are not Foucauldian or Habermasian, yet are frequently cited along with both, and therefore required for an accurate analysis of their intersection.

The criteria for inclusion of candidate sources were (a) a sustained theoretical engagement with ideas from Foucault and/or Habermas, rather than just passing references, (b) coverage of issues around artificial intelligence, algorithms, or similar technologies, rather than general surveillance or communication technology, and (c) connection to issues of democratic politics and governance. Excluded were those that referred to Foucault or Habermas only briefly, those that treated artificial intelligence only peripherally (such as an illustration in an argument about general digital media issues), and those that did not appear in peer-reviewed journals. Opinion pieces and unrefereed preprints would have been excluded except for the rare case of a uniquely significant theoretical contribution by a respected scholar that was not yet peer reviewed.

The literature search covered articles from peer-reviewed journals, book chapters, and scholarly books written in the English language and published from 2012 to 2026. The selection of this 15-year period has been intentional as it encompasses two overlapping bodies of literature, one dating back to 2012-2019 concerned with the reinterpretation of Foucauldian and Habermasian theories for the general context of digital and networked surveillance, and another one published from 2022-2026 concerning generative AI, large language models, and AI-based deliberation platforms. Key primary texts written by Foucault and Habermas that preceded this period have been included due to their theoretical contribution.

The sources have been identified by conducting structured searches of Google Scholar, database of journals indexed in Scopus, JSTOR, SAGE Journals, Taylor and Francis Online, SpringerLink, Phil Papers, and Oxford Academic employing combinations of the search terms "Foucault," "panopticon," "biopolitics," "governmentality," "surveillance," "Habermas," "public sphere," "deliberative democracy," "communicative action," and "artificial intelligence." The first search generated some 100 records of sources that could serve as candidates to either of the two theoretical strands. This set of 100 candidate records was filtered out in two rounds. First, the records were analyzed on the basis of the exclusion criteria outlined above. Second, the remaining 100 records were scrutinized on the basis of the three inclusion criteria outlined above, which resulted in filtering out 80 sources and selecting

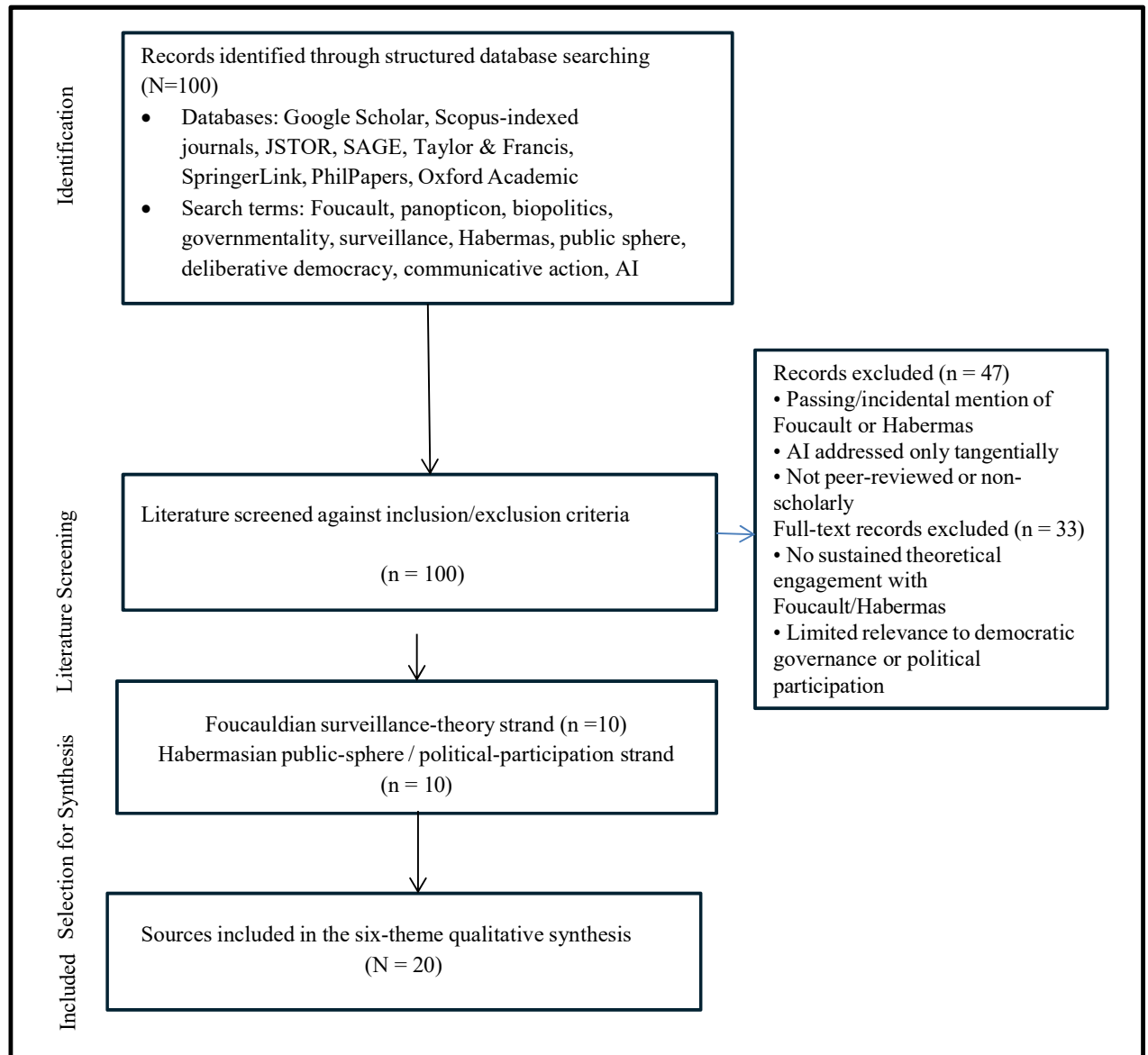
20 core sources for further analysis, ten in regard to the Foucauldian surveillance theory, and ten - regarding Habermasian theory of political participation and deliberation. The process is illustrated by Figure 1 below.

Figure 1

PRISMA-Style Flow Diagram of Source Identification and Selection

Figure 1

PRISMA-style flow diagram of the article identification, screening, and selection process (n = 100 identified; n = 20 included).



Framework for Analysis and Synthesis

The 20 core sources were read thematically. Each source was coded for its central

theoretical claims, its empirical or illustrative referents (such as predictive policing, social credit systems, algorithmic content moderation, or AI-mediated deliberation platforms), and its normative conclusions about democracy. Recurring codes were grouped inductively into six cross-cutting themes, reported in the Discussion section below, that capture both convergences and productive tensions between the Foucauldian and Habermasian literatures. To make the distribution of sources across themes transparent and auditable, a source-by-theme matrix is provided in Appendix A, allowing the reader to trace each thematic claim made in the Discussion section back to the specific sources that support it.

Credibility and Trustworthiness

Several steps were taken to strengthen the credibility of the synthesis. First, only peer-reviewed or otherwise clearly vetted scholarly sources were included in the 20-source core, and every core citation was independently verified against its original publication before inclusion. Second, each of the six themes reported in the Discussion section is supported by multiple, independently authored sources rather than by a single source, a form of source triangulation intended to guard against overreliance on any one author's idiosyncratic reading of Foucault or Habermas. Third, the coding and thematic grouping process was conducted iteratively, with initial themes revisited and refined as additional sources were read, consistent with recognized practice in qualitative and interpretive literature synthesis. A limitation should nonetheless be acknowledged: because the two theoretical traditions have developed largely in parallel, with limited direct cross-citation between them, most reviewed sources do not themselves attempt the kind of integration performed here. The synthetic claims presented in the Discussion section are therefore the interpretive contribution of this paper rather than a straightforward summary of a pre-existing consensus in the literature, and readers should weigh them accordingly as a theoretical argument rather than an empirical finding.

Discussion: Emergent Themes

Reading the Foucauldian and Habermasian literatures together produces six cross-cutting themes. Each is discussed in turn below, with explicit attention to where the two traditions converge, diverge, and productively complicate one another.

From the Panopticon to Algorithmic Governmentality

The most consistent finding across the Foucauldian literature is that Bentham's panopticon, as reconstructed by Foucault (1977), remains a generative but insufficient metaphor for AI-enabled surveillance. Where the panopticon relied on the visibility of a central watchtower and the internalized uncertainty of being watched, contemporary AI systems distribute surveillance across a diffuse, largely invisible infrastructure of sensors, data brokers, and predictive models (Dai et al., 2025; Galič et al., 2017). Several sources argue that this shift is better captured by Foucault's later concept of governmentality, the "conduct of conduct" through which populations are governed not by direct coercion but by shaping the field of

possible action (Gane, 2012; Varela, 2024). Predictive policing, China's social credit system, and algorithmic risk-scoring in the criminal justice and welfare systems are repeatedly cited as paradigm cases in which AI functions less as a watchtower and more as an ambient, self-adjusting apparatus of normalization (Sahakyan et al., 2025; Varela, 2024).

A related strand of this literature emphasizes that the disciplinary work once performed by an external observer is increasingly performed by the observed themselves. Because citizens cannot be certain when, or by what means, an algorithmic system is processing their data, many internalize a generalized caution that shapes behavior even absent any specific act of monitoring, a dynamic Galic et al. (2017) trace across successive waves of surveillance theory from the architectural panopticon to networked, infrastructural, and participatory forms of watching. This internalization is what several sources describe as a post-panoptic condition: subjects are disciplined not merely by the knowledge that they might be watched but by the predictive feedback loops through which algorithmic systems continuously adjust the information, opportunities, and prices presented to them (Dai et al., 2025). The Habermasian literature does not use the vocabulary of governmentality, but it registers a structurally similar concern in different terms: Jungherr and Schroeder (2023) argue that AI-enabled intermediary platforms increasingly determine what becomes visible within the public arena and what is pushed out of sight, a gatekeeping function that, while not disciplinary in the strict Foucauldian sense, performs an analogous work of shaping the field of public visibility within which citizens can act. Reading these two accounts side by side suggests that algorithmic governmentality operates simultaneously at the level of individual conduct, through predictive personalization, and at the level of collective visibility, through platform-level curation, a dual operation that neither tradition alone fully names.

Bio-politic and the Datafied Citizen

A second theme concerns the biopolitical management of populations through data aggregation. Foucault's concept of biopower, the administration of life at the level of population rather than the individual body, is repeatedly invoked to describe how AI systems classify, score, and manage citizens as statistical aggregates (Capodivacca & Giacomini, 2024; Varela, 2024). This literature emphasizes that AI's biopolitical function operates prior to, and often beneath, any single act of surveillance: it is the ongoing construction of predictive categories, risk profiles, and behavioral norms that most powerfully shapes citizens' life chances. Harcourt's (2018) account of the shift from the panopticon to an "expository society," in which subjects increasingly expose themselves voluntarily through digital participation, extends this concern to the question of consent, suggesting that biopolitical control today operates as much through the cultivation of self-disclosure as through coercive observation. This theme has no direct equivalent in the Habermasian corpus reviewed here, which tends to treat citizens as discursive agents capable of rational deliberation rather than as

statistical populations to be managed; this is one of the clearest points of theoretical divergence between the two traditions, and one to which the paper returns in the Conclusion.

Surveillance Capitalism and the Colonization of the Lifeworld

A third theme functions as a bridge between the two traditions. Zuboff's (2019) theory of surveillance capitalism, while not itself Foucauldian, is repeatedly cited alongside Foucault in the reviewed literature as an account of how commercial AI systems extract behavioral data as a raw material for prediction and profit (Gane, 2012). This economic reading of surveillance resonates strongly with Habermas's (1996) concern that the "system" organized around the media of money and administrative power, colonizes the "lifeworld," the taken-for-granted horizon of shared meaning in which communicative action takes place. Several Habermasian sources argue explicitly that AI-driven platforms accelerate this colonization by converting communicative exchanges, ordinary posts, searches, and conversations, into data commodities, thereby subordinating the norms of mutual understanding to the imperatives of engagement-maximizing algorithms (Jungherr & Schroeder, 2023). Read together, the Foucauldian and Habermasian literatures suggest that surveillance capitalism is simultaneously a disciplinary apparatus, in the Foucauldian sense of shaping conduct through data-driven normalization, and a colonizing force, in the Habermasian sense of displacing communicative rationality with instrumental and strategic rationality.

The Algorithmic Restructuring of the Public Sphere

The fourth theme, drawn primarily from the Habermasian literature, concerns the structural transformation of the public sphere under conditions of algorithmic mediation.

Habermas (2022) himself has argued that digital platforms have produced a further structural transformation beyond the one he described in his classic 1962 study, marked by the fragmentation of a once-unified bourgeois public into a proliferation of decentralized, often mutually unintelligible digital publics. Jungherr and Schroeder (2023) extend this argument empirically, contending that AI-enabled content curation and generation both fragment the public arena into algorithmically sorted enclaves and, paradoxically, threaten to re-concentrate control in the hands of a small number of platform gatekeepers whose algorithms determine what becomes collectively visible. This theme intersects with the Foucauldian literature at the point of gatekeeping power: where Habermasian scholars worry that algorithmic curation undermines the normative conditions for rational-critical debate, Foucauldian scholars would describe the same curation as a disciplinary technology that trains citizens, through the repeated reinforcement of engagement-optimized content, to attend to some issues and ignore others (Dai et al., 2025). The two vocabularies thus converge on a common empirical referent, algorithmic curation, while offering different diagnostic emphases: one normative and communicative, the other disciplinary and behavioral.

Communicative Action Machines and the Limits of AI-Mediated Deliberation

The fifth theme is the most novel in the recent literature: the emergence of AI systems explicitly designed to facilitate, rather than merely mediate or curate, democratic deliberation. Tessler et al. (2024) report that an AI mediator, explicitly named the "Habermas Machine" after Jürgen Habermas's theory of communicative action, was able to generate group statements that garnered higher endorsement from human discussants than statements produced by human mediators, while leaving groups less divided after deliberation. This finding has generated a lively secondary literature assessing whether such systems genuinely approximate Habermas's ideal speech situation, in which the strongest argument prevails free of coercion, or whether they merely simulate consensus without satisfying its underlying normative conditions (Volpe, 2025). Several sources caution that Habermas's own theory does not straightforwardly endorse the substitution of an algorithmic aggregator for embodied, reciprocal argumentation among citizens, since the legitimacy of communicative action depends on participants' mutual recognition of one another as accountable interlocutors, a condition an AI mediator can simulate but arguably cannot fully satisfy (Innerarity, 2024). Read through a Foucauldian lens, the Habermas Machine also raises a disciplinary question largely absent from the deliberative-democracy literature itself: an AI system trained to produce maximally agreeable group statements is, among other things, a technology that learns to predict and shape what citizens will find acceptable, a function structurally continuous with the normalizing operations Foucault (1977) associated with disciplinary power, even when its stated purpose is emancipatory rather than repressive.

Resistance, Accountability, and Democratic Repair

The sixth and final theme concerns the prospects for resisting or repairing the democratic harms identified in the preceding themes. The Foucauldian literature emphasizes counter-disciplinary strategies: export controls on surveillance technology, investigative exposure by civil-society organizations, and strategic litigation aimed at constraining abusive uses of AI surveillance, framed as ways of allowing citizens to "refuse the role of docile bodies" (Dai et al., 2025; Sahakyan et al., 2025). The Habermasian literature emphasizes institutional and design-based repair: strengthening independent public-service media, redesigning platform architectures to privilege reasoned exchange over engagement metrics, and, more speculatively, developing AI deliberation tools whose design is explicitly answerable to Habermasian norms of inclusion, sincerity, and non-coercion (Innerarity, 2023; Lafont, 2020). Both literatures converge on a shared, if differently justified, conclusion: neither surveillance nor deliberation is technologically determined, and the democratic character of AI systems depends substantially on the legal, institutional, and design choices that govern their deployment.

Areas of Convergence and Divergence

The two traditions converge on a shared diagnosis: AI-driven infrastructures concentrate power in opaque, poorly accountable hands, whether platform gatekeepers or state agencies, and both call for structural, not merely individual, remedies (Dai et al., 2025; Jungherr & Schroeder, 2023). They diverge sharply in unit of analysis and normative orientation. Foucauldian theory treats citizens as datafied populations shaped by disciplinary and biopolitical mechanisms operating beneath conscious awareness, and is comparatively skeptical of emancipatory claims (Varela, 2024). Habermasian theory treats citizens as reasoning agents capable of rational-critical discourse, and remains cautiously optimistic that AI can be designed to support communicative action (Tessler et al., 2024). Foucault emphasizes resistance through exposure and litigation; Habermas emphasizes repair through institutional and deliberative design (Innerarity, 2023; Sahakyan et al., 2025).

Conclusion

This synthesis suggests that artificial intelligence's relationship to democracy cannot be adequately captured by either Foucauldian surveillance theory or Habermasian public-sphere theory alone. Foucault's vocabulary of discipline, biopolitics, and governmentality illuminates how AI systems watch, classify, and normalize citizens, often beneath the threshold of conscious awareness, extending the logic of the panopticon into an ambient, algorithmically distributed form of governmentality. Habermas's vocabulary of the public sphere and communicative action illuminates how AI systems structure the discursive conditions under which citizens form opinions and hold power accountable, both by fragmenting and re-concentrating the public arena and, in its most recent and provocative applications, by attempting to engineer consensus directly. The six themes developed above indicate that these two dynamics are not parallel but entangled: the same predictive infrastructures that discipline citizens as data subjects also determine what becomes visible, discussable, and arguable in the public sphere; and the same deliberative technologies designed to realize Habermasian ideals of rational consensus are also, inescapably, technologies of behavioral prediction and normalization in the Foucauldian sense. AI, on this synthesis, does not simply strengthen or simply threaten democracy. It reconfigures the terms on which both surveillance and deliberation occur, often within a single technological system.

This entanglement also clarifies why debates that treat AI's democratic impact as a single, unified question, asking simply whether AI is "good" or "bad" for democracy, tend to generate more heat than light. The synthesis developed here suggests instead that the relevant question is always relational: good, or harmful, for which democratic function, assessed against which normative standard, and subject to which institutional safeguards? A facial-recognition system deployed for crowd control at a protest and a large language model deployed to summarize citizen submissions to a public consultation may share underlying

technical architectures, yet they engage almost entirely different democratic stakes, the former primarily disciplinary and biopolitical, the latter primarily communicative and deliberative, even though, as Theme 5 shows, no deliberative application is ever fully free of a disciplinary residue. A combined Foucauldian-Habermasian framework offers researchers and policymakers alike a more precise vocabulary for locating exactly where, along this spectrum, a given AI application sits.

References

- Capodivacca, S., & Giacomini, G. (2024). Discipline and power in the digital age: Critical reflections from Foucault's thought. *Foucault Studies*, 36, 227–251.
<https://doi.org/10.22439/fs.i36.7215>
- Dai, R., Thomas, M. K. E., & Rawolle, S. (2025). Revisiting Foucault's panopticon: How does AI surveillance transform educational norms? *British Journal of Sociology of Education*, 46(5), 650–668. <https://doi.org/10.1080/01425692.2025.2501118>
- Foucault, M. (1977). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). Vintage Books.
- Foucault, M. (2008). *The birth of biopolitics: Lectures at the Collège de France, 1978–1979* (G. Burchell, Trans.). Palgrave Macmillan.
- Galič, M., Timan, T., & Koops, B.-J. (2017). Bentham, Deleuze and beyond: An overview of surveillance theories from the panopticon to participation. *Philosophy & Technology*, 30(1), 9–37. <https://doi.org/10.1007/s13347-016-0219-1>
- Gane, N. (2012). The governmentalities of neoliberalism: Panopticism, post-panopticism and beyond. *The Sociological Review*, 60(4), 611–634. <https://doi.org/10.1111/j.1467-954X.2012.02126.x>
- Habermas, J. (1991). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society* (T. Burger, Trans.). MIT Press. (Original work published 1962)
- Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy* (W. Rehg, Trans.). MIT Press.
- Habermas, J. (2022). Reflections and hypotheses on a further structural transformation of the political public sphere. *Theory, Culture & Society*, 39(4), 145–171.
<https://doi.org/10.1177/02632764221112341>
- Harcourt, B. E. (2018). Virtual transparency: From the panopticon to the expository society and beyond. In B. E. Harcourt, *Transparency, society and subjectivity* (pp. 369–391). Springer International Publishing. https://doi.org/10.1007/978-3-319-77161-8_18
- Innerarity, D. (2023). Predicting the past: A philosophical critique of predictive analytics. *IDP. Revista d'Internet, Dret i Política*, 39, 1–12.
- Innerarity, D. (2024). The epistemic impossibility of an artificial intelligence take-over of

- democracy. *AI & Society*, 39(4), 1667–1671. <https://doi.org/10.1007/s00146-023-01631-y>
- Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social Media + Society*, 9(3), 1–14. <https://doi.org/10.1177/20563051231186353>
- Jungherr, A., & Schroeder, R. (2023). Artificial intelligence and the public arena. *Communication Theory*, 33(2–3), 164–173. <https://doi.org/10.1093/ct/qtad006>
- Lafont, C. (2020). *Democracy without shortcuts: A participatory conception of deliberative democracy*. Oxford University Press.
- Sahakyan, H., Gevorgyan, A., & Malkjyan, A. (2025). From disciplinary societies to algorithmic control: Rethinking Foucault's human subject in the digital age. *Philosophies*, 10(4), Article 73. <https://doi.org/10.3390/philosophies10040073>
- Tessler, M. H., Bakker, M. A., Jarrett, D., Sheahan, H., Chadwick, M. J., Koster, R., Evans, G., Campbell-Gillingham, L., Collins, T., Parkes, D. C., Botvinick, M., & Summerfield, C. (2024). AI can help humans find common ground in democratic deliberation. *Science*, 386(6719), eadq2852. <https://doi.org/10.1126/science.adq2852>
- Varela, D. T. (2024). Artificial intelligence law through the lens of Michel Foucault: Biopower, surveillance, and the reconfiguration of legal normativity. *Open Journal of Social Sciences*, 12(12), 189–201. <https://doi.org/10.4236/jss.2024.1212014>
- Volpe, A. (2025). Toward an artificial deliberation? On Google DeepMind's Habermas Machine. *Ethics and Information Technology*, 27(4), Article 45. <https://doi.org/10.1007/s10676-025-09854-1>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the frontier of power*. PublicAffairs.