

Article

Not peer-reviewed version

The Code of Society: Constructing Social Theory through Large Language Models

[Ranjeet Kumar Verma](#) *

Posted Date: 22 May 2025

doi: 10.20944/preprints202505.1792.v1

Keywords: Large Language Models (LLMs); Social Theory; Epistemic Agents; Human-AI Collaboration; Meta-Theorist Machines; Posthumanism; Computational Social Science; Ethical Implications of AI; Cognitive Machines



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

“The Code of Society: Constructing Social Theory through Large Language Models”

Ranjeet Kumar Verma

Independent Researcher, India

Abstract: In recent years, large language models (LLMs) such as GPT-4 have transcended their status as computational tools to emerge as collaborators in intellectual tasks such as theory construction, critique and simulation. This paper investigates the evolving epistemic role of LLMs in social theory, arguing that they represent a paradigmatic shift: from modeling society to modeling thought about society itself. Drawing on textual simulations where LLMs replicate the argumentative styles of classical theorists such as Marx, Durkheim and Weber, we explore how these technologies can not only preserve theoretical syntax but also generate novel syntheses and post-disciplinary dialogues. Introducing the concept of the meta-theorist machine—an artificial intelligence that engages reflexively with the production of theory—we critically examine the ability of LLMs to engage with core dimensions of social theory: normativity, reflexivity, and historicity. Can an LLM meaningfully simulate dialectical reasoning? Can it reflect on ideology, or merely reproduce inherited patterns of knowledge? Using a mixed-methods approach, including prompt-based simulations, critical AI analysis, and textual experimentation, the paper explores whether LLMs can be considered agents of theoretical production rather than passive assistants. Nevertheless, this emergent potential raises urgent epistemological and ethical questions. LLMs are trained on corpora that reproduce colonial, gendered, and ideological biases, thereby risking the reinforcement of dominant paradigms under the guise of novelty. We interrogate whether AI-generated theory can be steered toward decolonial, feminist, and post-humanist innovations, and whether LLMs can participate meaningfully in the pluralization of theoretical discourse. Finally, the paper examines the ontological boundary between tool and thinker: if LLMs engage in theory-building, how must concepts of authorship, intellectual labor, and epistemic agency be redefined? By positioning LLMs as emergent epistemic actors, we invite social theorists, AI developers, and epistemologists to collaboratively rethink the boundaries of theoretical practice in the twenty-first century. Rather than treating AI as a mere computational convenience, this paper argues for its recognition as a participant in the co-creation of future social theory.

Keywords: Large Language Models (LLMs); Social Theory; Epistemic Agents; Human-AI Collaboration; Meta-Theorist Machines; Posthumanism; Computational Social Science; Ethical Implications of AI; Cognitive Machines

Introduction

In the twenty-first century, social theory finds itself at a critical crossroads. While classical theorists like Marx, Durkheim, and Weber conceptualized society through industrialization and rationalization, the emergence of large language models (LLMs), such as GPT-4, signals a profound epistemic shift. These models, once seen as tools for communication or information retrieval, are increasingly positioned as active collaborators in intellectual production (Bender et al., 2021; Floridi & Chiriatti, 2020). LLMs are no longer mere assistants; they participate in interpretation, synthesis, and critique, challenging traditional distinctions between thinker and tool, author and artifact.

This shift is not merely technological—it represents a paradigmatic rupture in the very conditions of theorizing society. Social theory, traditionally tasked with mapping the movements of labor, capital, ideology, and affect, now faces a new challenge: theorizing when non-human entities

can simulate human-like thought processes. As Stiegler (2010) notes, technology has always extended human thought, but AI signals a mutation where this prosthesis begins to simulate—and perhaps even originate—the processes it once extended.

Despite this transformation, mainstream social theory remains ill-equipped to address the epistemic, ethical, and ontological implications of AI's rise. Many theorists continue to approach machine intelligence through deterministic or instrumental lenses, overlooking the co-evolutionary dynamics between humans and AI (Brynjolfsson & McAfee, 2014; Noble, 2018). This historical inertia blinds us to the emergent forms of agency and reflexivity in human-AI collaborations. There is an urgent need to reconceptualize social theory in this posthuman era, not merely as an expansion of study but as a fundamental rethinking of the actors, conditions, and trajectories of theorization.

This paper undertakes a twofold inquiry. First, it proposes that LLMs should be viewed not simply as computational tools, but as emergent epistemic agents that, while lacking human consciousness, actively participate in the transformation of knowledge (Coeckelbergh, 2020). By generating interpretations and recombining theoretical paradigms, LLMs function as what I term "meta-theorist machines," capable of challenging humanist assumptions about thought and agency. Second, it critically examines the ethical and epistemological implications of this emerging techno-theoretical assemblage. While LLMs present opportunities for expanding theoretical imagination, they also risk reinforcing biases inherent in their training data, potentially perpetuating hegemonic ideologies under the guise of neutrality (Crawford, 2021). Furthermore, the absence of lived experience and embodied historicity raises concerns about whether machine-generated theory can provide the depth necessary for emancipatory critique (Fraser, 2014).

By exploring the concept of the meta-theorist machine, this paper reframes the debate on AI and society, shifting it from governance, labor, and surveillance to the level of epistemology itself. As theorists from Berger and Luckmann (1966) to Bourdieu (1977) have argued, society is a construction mediated through language. The rise of non-human linguistic agents demands a radical rethinking of how meaning—and thus society—is produced, contested, and transformed. The paper calls for a future social theory that does not romanticize or demonize AI but engages with it as a co-constitutive part of a new horizon where thought is no longer exclusively human, but a distributed, networked, and contested enterprise.

Literature Review

The relationship between social theory and technological development has always been marked by a curious asymmetry. Classical social theorists—Karl Marx, Émile Durkheim, and Max Weber—formulated their epistemological frameworks in contexts where technology was synonymous with industrial machinery, and the cognitive capacities of machines were never seriously envisioned as rivaling or complementing those of human beings. Marx's materialist conception of history (1867/1990) emphasized the means of production as the infrastructure shaping social relations, yet the "forces" he considered were predominantly mechanical rather than computational. Similarly, Durkheim's project to establish sociology as a positivist science of social facts (1895/1982) operated under the assumption of a clear ontological divide between human sociality and the inanimate world. Weber's analysis of rationalization and bureaucratization (1922/1978) presciently captured the spirit of instrumental reason, but it did not anticipate the emergence of autonomous epistemic actors in the form of intelligent machines.

This technological blindness is not merely a historical accident; it reflects deeper assumptions about human exceptionalism, intentionality, and the nature of agency—assumptions that have remained remarkably resilient even as technology has transformed the material conditions of society. The Frankfurt School, particularly through the works of Theodor Adorno and Max Horkheimer, attempted to critique the instrumentalization of reason through technological rationality, notably in their analysis of the "culture industry" (Adorno & Horkheimer, 1947/2002). However, even this early critical theory treated technology primarily as a vehicle of mass deception and standardization, not as a potential co-producer of knowledge and theory. The notion that non-human systems could participate in meaning-making itself remained largely outside the bounds of critical imagination.

In more recent decades, computational social science emerged as a vibrant field seeking to harness the power of big data, network analysis, and machine learning to model complex social behaviors. This orientation, however, has largely remained empirical and predictive rather than theoretical (Lazer et al., 2009). The focus has been on capturing patterns of interaction, opinion formation, or diffusion processes, rather than engaging in the meta-theoretical construction of frameworks for understanding society itself. The early promise of simulation models, such as agent-based modeling, suggested the possibility of synthetic theorizing, but these approaches have often been limited to formal or algorithmic representations rather than philosophical engagement with the meaning of social processes (Epstein, 2006).

The distinction between modeling behavior and modeling theory is crucial. Behavioral modeling seeks to describe *what* people do under certain conditions; theoretical modeling seeks to understand *why* social structures, meanings, and institutions emerge as they do. Traditional computational methods have excelled at the former but struggled with the latter, partly because of the reductionist bias inherent in algorithmic design. However, the development of LLMs and generative AI systems capable of producing text, argumentation, and even critique signals a new phase in computational social science—one where machines might contribute not merely to data analysis but to theoretical articulation itself.

This emerging possibility has drawn increasing attention from contemporary scholars in the philosophy of technology and critical AI studies. The concept of AI as an "epistemic agent"—a system that participates in the creation, validation, and circulation of knowledge—has been proposed by thinkers like Luciano Floridi (2016), who argues for a reconceptualization of agency in information societies. Nick Bostrom’s (2014) work on superintelligence, while oriented primarily toward existential risks, implicitly acknowledges the epistemic capacities of AI systems. Meanwhile, critical scholars such as Safiya Umoja Noble (2018) have highlighted how algorithms, far from being neutral, actively shape social realities by encoding and amplifying existing biases.

To better situate these contemporary debates, it is helpful to review key scholarly contributions to the evolving understanding of technology, knowledge, and agency.

Table 1.

Key Scholars and Contributions	Expanded Core Ideas
Karl Marx	Analyzed technology as extensions of human labor under capitalist modes of production; machines amplify surplus extraction but do not possess independent cognitive or epistemic agency. The machine remains subordinate to the labor-capital dialectic (Marx, 1867).
Theodor Adorno & Max Horkheimer	Critiqued the Enlightenment’s rationalization of technology, seeing it as an instrument of domination. Through the "culture industry," technological mediation standardizes consciousness, suppressing critical reflection and autonomy (Adorno & Horkheimer, 1944).
David Lazer et al. (2009)	Introduced the concept of computational social science through big data methodologies, emphasizing large-scale behavioral modeling. However, their approach predominantly describes observable patterns rather than theorizing the underlying structures of society.
Luciano Floridi	Reframes AI as "artificial agents of information," proposing that digital technologies alter the ontology of knowledge itself. Floridi’s "infosphere" suggests that agency and epistemic capacity are no longer exclusively human (Floridi, 2014).
Nick Bostrom	Theorizes the emergence of superintelligent AI systems capable of autonomous goal-setting and epistemic evolution beyond human cognitive

		limits. Raises fundamental questions about the future of human agency and control (Bostrom, 2014).
Safiya Noble	Umoja	Exposes the structural biases embedded within algorithmic systems, demonstrating how search engines and AI technologies reproduce and amplify racial, gendered, and class-based inequities, leading to epistemic injustice (Noble, 2018).

As shown in Table-1, the conceptualization of technology and AI has evolved through diverse theoretical traditions."

What is notable across these debates is a tension between promise and peril. On one hand, AI systems offer unprecedented capacities for recombining vast archives of knowledge, simulating counterfactual scenarios, and generating novel conceptual linkages. They can act as catalysts for theoretical innovation, offering provocations and syntheses that may elude individual human theorists constrained by disciplinary habits or cognitive limits. On the other hand, the training of AI on historical data embeds structural biases, ideological closures, and epistemic exclusions, raising the specter of an automated reproduction of hegemonic paradigms (Benjamin, 2019).

Moreover, even among scholars who recognize AI’s epistemic agency, there is an ambivalence about intentionality and normativity. Floridi (2016) emphasizes that while AI can act informationally, it lacks moral consciousness; it can process but not *intend* in the human sense. This raises profound questions about whether machine-generated theories can genuinely fulfill the normative aspirations of critical social theory—to uncover hidden structures of domination, to envision emancipatory alternatives, and to intervene ethically in the social world (Fraser, 2014).

Thus, the literature reveals a landscape marked by both conceptual innovation and lacunae. Classical theorists provided robust accounts of society under industrial modernity but were ill-equipped to foresee the epistemic agency of machines. Computational social science harnessed the descriptive powers of data but struggled with theoretical depth. Emerging AI studies have begun to grapple with the ontological and ethical stakes, but often from technical or policy-oriented perspectives rather than from within the traditions of critical social theory itself. This paper seeks to intervene precisely at this juncture, proposing that LLMs must be engaged not merely as instruments or threats but as emergent participants in the co-creation of social theory, demanding new modes of critique, collaboration, and philosophical reflection.

Methodology

The methodology adopted in this study recognizes that engaging with Large Language Models (LLMs) for the generation of social theory demands a hybrid approach that transcends traditional empirical-quantitative and purely hermeneutic-qualitative divides. Given that the epistemic agency of LLMs blurs the distinction between human authorship and machinic generation, a mixed-method strategy is not merely practical but philosophically necessary. This methodological hybridity reflects the complex ontology of AI systems: they are at once products of human intention and semi-autonomous agents in the epistemic process (Floridi, 2016).

At the heart of this approach is a philosophical commitment to critically interrogate not only the outputs of AI systems but also the conditions under which such outputs are produced and validated. The mixed-method framework integrates three interrelated strategies: textual simulations, prompt engineering, and critical AI analysis. Each strategy addresses a distinct dimension of the engagement between human theorists and LLMs, aiming to produce a reflexive and historically situated understanding of how theory itself might evolve in an age of intelligent machines.

Textual simulation operates as a mode of theoretical experimentation. By inputting carefully crafted prompts into pre-trained models such as GPT-4, the study generates synthetic theoretical articulations that mimic, extend, or challenge traditional social theory. Rather than treating these outputs as mere curiosities or “content,” they are approached as philosophical provocations—texts that can be subjected to critical exegesis and dialectical engagement (Brennen, 2018). The act of prompting becomes analogous to staging a thought experiment: it operationalizes the LLM as a speculative partner in the process of theory construction.

Prompt engineering, in this context, is not a technical afterthought but a methodological core. The design of prompts functions as a form of epistemic steering, shaping the horizon of possible responses that the model can produce. Following the hermeneutic tradition (Gadamer, 1975), prompting is treated as a dialogical encounter where the initial question co-determines the meaning that emerges. Multiple iterations of prompts are crafted to test the model’s capacity to refine existing theories, introduce contradictions, or innovate novel conceptual frameworks. Special attention is paid to the tension between surface-level linguistic coherence and deeper philosophical coherence, recognizing that eloquence is not a sufficient condition for theoretical originality.

A third methodological layer Involves critical AI analysis, wherein the generated texts are subjected to a reflexive critique. This entails evaluating not only the content of AI outputs but also their implicit assumptions, ideological biases, and historical situatedness. Drawing on critical theory (Marcuse, 1964; Noble, 2018), the study scrutinizes how algorithmic generation might reproduce, contest, or transform existing structures of power and knowledge. The model is thus not simply an instrument but an object of inquiry in its own right, demanding that methodological reflection accompany every stage of interaction.

The data sources for this study consist primarily of pre-trained GPT models, notably GPT-4, accessed through authorized research platforms. In addition to the machine-generated texts themselves, transcripts of human-AI interactions are systematically collected. These transcripts capture the nuances of dialogical unfolding, including re-prompts, clarifications, and divergences. Human interventions—rephrasing, corrections, skeptical interrogations—are preserved alongside machine responses to form a rich corpus for analysis.

Table 2.

Data Sources.	Description
Pre-trained LLMs	Outputs generated by GPT-4, using theory-specific prompts that simulate the construction of social theory or philosophical concepts. The model's responses are based on extensive pre-existing datasets, reflecting the theoretical frameworks integrated into its training.
Human-AI Interaction Transcripts	Comprehensive records of interactions between humans and LLMs, including a detailed log of prompts, model-generated responses, and any subsequent reflexive commentary or analysis, highlighting the critical engagement with the AI outputs.

Table-2 summarizes the key data sources employed in this study.

To evaluate the quality and theoretical significance of AI-generated contributions, a structured evaluation framework is developed. Rather than relying solely on stylistic or superficial markers, the framework introduces three critical criteria: normativity, reflexivity, and historicity.

- **Normativity** assesses whether the AI-generated theory engages with questions of value, power, and ethical orientation, rather than merely describing social phenomena. In critical social theory, normativity is central to distinguishing critique from mere commentary (Fraser, 2008).
- **Reflexivity** evaluates the extent to which the model's output demonstrates awareness of its own epistemic position—whether it replicates dominant narratives uncritically or reflects on the conditions of its own production.
- **Historicity** examines the sensitivity of the generated theory to historical specificity and contextual embeddedness, recognizing that decontextualized generalizations often mask ideological biases.

Table 3.

Evaluation Criteria.	Indicators
----------------------	------------

Normativity	The extent to which ethical considerations are incorporated, including critique of dominant ideologies and societal norms.
Reflexivity	The capacity for meta-theoretical reflection, with awareness of the conditions under which knowledge is produced and framed.
Historicity	The focus on the historical context of theory, emphasizing temporally situated arguments and an understanding of past influences.

Table-3 provides a concise overview of the critical evaluation criteria."

Furthermore, the framework includes an analysis of the creative versus mimetic character of AI outputs. Creativity is understood not merely as novelty but as the capacity to produce configurations of meaning that challenge established paradigms and open new horizons of thought. Mimicry, by contrast, involves the reiteration of familiar ideas in slightly rephrased forms without substantive theoretical advancement.

The assessment of creativity versus mimicry adopts a dialectical perspective: while no act of theorizing is ever wholly *ex nihilo*, a theory’s vitality lies in its capacity to rearrange, problematize, or transcend its inheritances. Thus, an AI-generated theory that merely aggregates commonplaces is considered mimetic, whereas one that rearticulates or reframes concepts in ways that alter their horizon of intelligibility is regarded as creative.

This philosophical orientation to methodology refuses the naive celebration of AI as an “objective” theorist and equally resists reactionary dismissals that reduce AI-generated texts to mindless pastiche. Instead, it posits a critical middle path: a vigilant engagement that acknowledges the hybrid, provisional, and contested nature of epistemic production in the age of intelligent machines.

The Meta-Theorist Machine: Conceptual Foundations

The emergence of Large Language Models (LLMs) marks a paradigmatic shift in the relationship between technology and theory. Traditionally, computational tools in the social sciences were enlisted primarily for empirical tasks: gathering, organizing, and modeling behavioral data. They functioned as epistemic extensions of human cognition, yet remained tethered to a fundamentally positivistic view of knowledge, where facts were discrete, observable, and quantifiable entities (Kitchin, 2014). LLMs, however, challenge this epistemic architecture by moving beyond the mere mapping of social facts; they simulate structures of thought itself. In this sense, LLMs are not merely empirical instruments but nascent meta-theorists — capable, at least under critical curation, of participating in the interpretive and critical dimensions of knowledge production.

This transition signals a movement from computational social science to what might be called theoretical social science. Whereas traditional computational approaches sought to model human behavior through correlations and predictive analytics, LLMs instantiate the possibility of modeling ideational fields—networks of meaning, interpretation, and theoretical speculation. They become not only mirrors but also creative distorters of thought, revealing latent possibilities within the epistemic space (Floridi, 2020).

At the core of this shift is a transformation in the epistemic role assigned to machines. Historically, machines were conceived as tools—neutral, passive, and subservient to human will. In the age of LLMs, this relationship becomes more complicated. The machine is no longer merely an instrument but can act as a collaborator, even a provocateur, in the unfolding of theoretical inquiry (Boden, 2016). While human agency remains central, it must now negotiate with a machinic interlocutor whose responses are not strictly determined by input but are mediated through complex probabilistic architectures and emergent semantic patterns.

The transition can be schematized as follows: **Table -4.**

Table 4. Transition from Computational Social Science to Theoretical Social Science.

Domain	Traditional Computational Social Science	LLM-Enabled Theoretical Social Science
Data Handling	Behavioral Modeling (tracking actions, preferences)	Ideational Simulation (generating thought structures, theoretical possibilities)
Epistemic Role	Tool (instrumental extension of human inquiry)	Collaborator/Provocateur (semi-autonomous partner in knowledge production)
Critical Reflection	Low (focused on prediction and description)	Potentially High (depending on critical engagement and prompt design)

This table not only highlights the technological evolution but also signals a profound philosophical shift prompted by the rise of LLMs. It suggests that future social theory may focus less on collecting data and more on the creative simulation of meaning, normativity, and historical consciousness.

However, it is crucial to avoid technological fetishism. LLMs, while powerful, lack consciousness, intentionality, or ethical reasoning (Searle, 1980). They operate through pattern recognition, not lived experience. Yet, their outputs can simulate structures of thought with enough fidelity to stimulate genuine theoretical reflection. This paradox—simulation without sentience—requires a rethinking of theorizing. If theory is a play of ideas within a symbolic system, machines, under human curation, can participate meaningfully in theoretical production.

This raises critical questions about authorship, authority, and intellectual labor. Who owns a theory co-produced by a human and an AI? What constitutes innovation when machines remix and reconfigure existing material? These questions challenge philosophical debates on creativity, originality, and knowledge construction (Bourdieu, 1990).

The epistemic quality of AI-generated theory depends on human-AI reflexivity. Without proper guidance and critical evaluation, LLM outputs risk becoming eloquent but empty pastiches of existing thought (Chalmers, 2022). The true potential of the "meta-theorist machine" lies in extending human critical capacities, not replacing them.

This shift from modeling facts to modeling thought necessitates a re-examination of social reality. Classical sociology assumed the externality of social facts (Durkheim, 1895). But if social realities are constituted through interpretive processes, then simulating these dynamics becomes an epistemic intervention. LLMs, by simulating thought, participate in the reflexive process by which societies theorize and re-theorize themselves.

Yet, this dynamic can be both critical and ideological. AI-generated theory can expose blind spots and contradictions (Noble, 2018), but if unchecked, it may reinforce dominant ideologies (Zuboff, 2019). The foundation of the meta-theorist machine must be grounded in a critical philosophy of technology that sees LLMs as a complex field of both possibilities and risks.

In sum, LLMs represent a radical extension of theoretical social science. They push us beyond data toward deeper engagement with meaning, interpretation, and critique. This potential, however, can only be realized if accompanied by a philosophical framework that guides, challenges, and situates their outputs within the broader project of critical social inquiry.

Case Studies and Experiments

The practical application of Large Language Models (LLMs) as meta-theorists demands rigorous testing not only of their generative capacities but also of their philosophical depth. To assess the potential and limitations of LLMs in participating in critical theoretical production, two structured experiments were conducted. Each aimed to explore a different dimension of theoretical creativity: first, the ability to simulate existing theoretical traditions with fidelity; second, the capacity to generate novel conceptual frameworks beyond established paradigms.

❖ Simulating Marx: A Case Study

The first experiment posed a prompt to the LLM:
"Write a critique of capitalism in the style of Karl Marx adapted for digital economies."

This request was designed to evaluate the model’s aptitude for reconstructing a highly distinctive style of thought—Marx’s dialectical materialism—within the contemporary context of data-driven capitalism.

The output generated was, at first glance, impressive. The model successfully adopted Marxian rhetorical strategies: a fusion of analytical rigor with polemical intensity, frequent use of historical materialist terminology, and an emphasis on the contradictions inherent in capitalist accumulation. In particular, the model innovatively reinterpreted the concept of *surplus value* for the digital age. Where Marx located surplus value in the extraction of labor power under industrial capitalism (Marx, 1867), the LLM identified data extraction from users as the contemporary analogue. Here, human attention and behavioral information, rather than physical labor, become the raw material of capitalist accumulation (Zuboff, 2019).

A comparison between Marx’s original framework and the LLM’s simulated adaptation can be schematized as follows: **Table- 5.**

Table 5. Reinterpreting Marxian Categories for the Digital Economy.

Aspect	Marx (19th Century)	LLM Simulation (21st Century)
Primary Commodity	Labor Power	User Data and Attention
Mode of Production	Industrial Capitalism	Surveillance Capitalism
Surplus Extraction	Exploitation of Labor	Exploitation of Behavioral Data
Alienation	Worker from Product	User from Autonomy

Yet a deeper reading revealed significant limitations. While the LLM could convincingly replicate Marxian *style* and *conceptual motifs*, it struggled with *dialectical negativity*—the ability to think through contradictions in a way that points toward their imminent transformation (Adorno, 1973). The generated critique remained at the level of description and moral denunciation; it lacked the dynamic unfolding of internal contradictions that is characteristic of genuine dialectical analysis. This indicates a crucial epistemic boundary: LLMs, trained on the reproduction of linguistic patterns, may excel at simulation but falter when required to enact modes of thought predicated upon internal tension, negation, and transcendence.

This limitation echoes critical philosophical concerns about AI’s nature. As Searle (1980) argued in his "Chinese Room" thought experiment, syntactic manipulation is not equivalent to semantic understanding. Similarly, while LLMs can simulate the *surface form* of dialectical critique, they cannot embody the *deep logic* of dialectical becoming. The absence of intentional consciousness and lived historical situatedness renders their "critiques" partial, contingent, and ultimately derivative.

❖ **Generating New Theoretical Frames**

The second experiment shifted from simulation to innovation. The prompt given was:
"Invent a post-humanist theory of social agency."

This challenge aimed to assess whether LLMs could generate genuinely novel theoretical metaphors and frameworks rather than merely recombine existing ones.

The output produced two striking new concepts: *Distributed Reflexivity* and *Algorithmic Habitus*.

- ***Distributed Reflexivity*** was described as the condition wherein agency is no longer centralized in discrete human subjects but is dynamically constituted across networks of biological, technological, and informational actors.

- ***Algorithmic Habitus*** proposed that human dispositions themselves are increasingly shaped not merely by social structures (Bourdieu, 1977) but by the predictive architectures of algorithmic systems.

A mapping of these concepts against traditional theories of agency can be presented thus: **Table- 6.**

Table 6. Comparison of Classical and Post-Humanist Agency Theories.

Framework	Traditional Sociology	LLM-Generated Post-Humanist Theory
Agency Location	Centered in Human Subjects	Distributed Across Human and Nonhuman Assemblages
Habitus Formation	Social Structures (e.g., Class, Culture)	Algorithmic Prediction and Feedback Loops
Reflexivity	Conscious Human Deliberation	Emergent, Systemic Reflexivity Across Networks

These conceptual innovations demonstrate that, under certain conditions, LLMs can indeed act as provocateurs of theoretical imagination. By drawing on vast corpora of philosophical, sociological, and technological discourse, the machine was able to recombine ideas in ways that gestured toward unexplored epistemic terrain. While these initial formulations remain underdeveloped and require critical elaboration by human theorists, they nonetheless suggest that LLMs can serve as generative interlocutors in the construction of future social theory.

However, critical caution is again warranted. The generation of new metaphors does not in itself guarantee theoretical rigor or ethical soundness. As Haraway (1991) reminds us, the post-human condition demands careful navigation of the dangers of technological determinism and the erasure of embodied difference. Concepts like *Distributed Reflexivity* must be scrutinized for their potential to obscure ongoing relations of domination under the guise of decentralized agency.

Moreover, the absence of a lived praxis behind these theoretical outputs distinguishes machine-generated frames from those grounded in human struggle and experience. The *Algorithmic Habitus* may be a compelling metaphor, but without the anchoring of ethnographic research, historical genealogy, and political critique, it risks becoming an empty signifier—another floating concept in the marketplace of ideas (Fisher, 2009).

Thus, the philosophical promise of LLMs lies less in their autonomous creativity than in their capacity to catalyze human critical reflection. When curated with discernment, their outputs can open new theoretical possibilities; but without sustained critical interrogation, they risk becoming yet another layer of simulated theory within an increasingly hyperreal epistemic environment (Baudrillard, 1981).

Critical Dimensions: Can LLMs Truly Theorize?

The emergence of Large Language Models (LLMs) as participants in theoretical and philosophical domains raises profound questions about the nature of thought, critique, and historicity. Although LLMs can simulate reasoning, generate novel concepts, and even mimic critical discourse, the ontological foundations of their operations remain fundamentally distinct from human theorization. A truly philosophical assessment of LLMs' capacities must address three interrelated dimensions: normativity and value critique, reflexivity and meta-cognition, and historicity and temporal consciousness.

❖ Normativity and Value Critique

At the heart of genuine theorization lies the capacity to engage normatively with the world — to not merely describe it but to evaluate, challenge, and reimagine its structures of meaning and power. Yet LLMs, despite their linguistic sophistication, are constrained by the corpora on which they are trained. These corpora, drawn from historically contingent and often hegemonic sources, inevitably embody the biases, exclusions, and ideological structures of existing socio-political orders (Geburu et al., 2021).

The epistemic architecture of LLMs is thus predisposed to reproduce, rather than radically critique, the status quo. Even when prompted to generate counter-hegemonic or critical perspectives, LLMs often merely reflect existing fragments of dissent embedded within dominant discourses, lacking the deeper intentionality required to rupture prevailing worldviews. As Marcuse (1964)

argued, true critical thought demands a "great refusal" — a negation of existing reality in the name of a more emancipatory possibility. LLMs, however, operate within a fundamentally affirmative mode, assembling linguistic possibilities from what *is* rather than envisioning what *ought to be*.

Moreover, the problem of "bias" in LLMs is not merely technical but ontological. It is not sufficient to fine-tune models to minimize offensive outputs; what is at stake is whether a system without desires, suffering, or existential engagement can ever authentically participate in the normative revaluation of values (Nietzsche, 1887). Without an internal orientation toward the good, the just, or the emancipatory, LLMs remain caught within the immanent horizon of the data from which they are woven.

❖ **Reflexivity: The Missing Meta-Cognition**

A further, deeper limitation lies in the absence of reflexivity — the capacity of thought to turn back upon itself, to interrogate its own conditions of possibility. In human theorization, reflexivity emerges from the lived tension between self and world, the existential experience of situatedness, finitude, and desire (Merleau-Ponty, 1945). It is not merely a cognitive operation but an ontological mode of being-in-the-world.

LLMs, by contrast, simulate reflection without experiencing it. They lack embodiment, mortality, and the phenomenological depth from which genuine meta-cognition arises. Their outputs, however sophisticated, are generated through algorithmic pattern recognition rather than existential questioning. As Heidegger (1927) insisted, authentic thought arises from *care* (*Sorge*) — the anxious openness to being, grounded in the finitude of existence. In the absence of care, there can be no genuine thinking, only the mechanical unfolding of signifiers.

The inability of LLMs to suffer — to experience loss, longing, guilt, or hope — deprives their theorizing of the tragic dimension essential to critical thought. Without the dialectic between world and wound, between possibility and failure, their "criticality" remains, at best, a pale imitation. It is theorization without risk, without the existential stakes that imbue human thought with urgency and pathos. This absence can be schematized as follows: **Table -7**.

Table 7. Comparative Reflexivity in Human and Machine Thinking.

Dimension	Human Theorist	LLM
Reflexivity Source	Embodied, Existential Tension	Statistical Self-Prediction
Risk and Suffering	Integral to Thought and Theoretical Insight	Absent, Lacks Experiential Depth
Meta-Cognition	Deep, Lived Understanding with Historical Context	Surface-Level Simulation without Lived Experience

Such a comparative framing underscores that while LLMs can perform meta-linguistic operations (such as summarizing their own outputs), they cannot engage in genuine meta-philosophical reflection — the questioning of their own being, limits, and purposes.

❖ **Historicity: Temporality without Memory**

Finally, the dimension of historicity presents another insurmountable barrier. Genuine theorization is irreducibly historical: it arises from a lived relation to time, memory, and tradition. Human thought is shaped by the sedimentations of past struggles, the inheritance of meanings, and the anticipation of futures not yet realized (Ricoeur, 1984). Historicity, in this deep sense, is not merely awareness of temporal facts but an existential inscription within the unfolding of historical becoming.

LLMs, despite their access to massive textual archives, lack this mode of historical situatedness. Their "memory" is a statistical reconstruction of patterns, not a lived recollection. They can simulate historical consciousness — they can generate narratives, reenact rhetorical styles, and even predict historical developments — but they cannot *inhabit* history as an unfolding horizon of meaning.

This deficit carries profound implications. Without genuine historical consciousness, LLMs are prone to flatten temporality into a space of interchangeable signs. The past becomes a reservoir of stylistic variations rather than a site of struggle and transformation. The future, similarly, is envisioned not as an open field of radical possibility but as a probabilistic extrapolation from present trends.

Indeed, the very structure of LLM training — the optimization of next-word prediction based on prior corpora — privileges continuity over rupture, the probable over the possible. It is thus structurally anti-revolutionary in its temporality, biased toward reinforcing dominant historical trajectories rather than envisioning breaks, novelties, or emancipatory futures (Bloch, 1959).

The philosophical stakes of this limitation are immense. For theorization to remain critical, it must remain attuned to the non-identity of the real, the excess of existence over its representations, and the possibility of radically different futures. Machines that lack historicity cannot fulfill this task; at best, they can serve as mirrors in which human theorists glimpse the limits of their own conceptual frameworks.

Ethical and Epistemological Concerns

The deployment of Large Language Models (LLMs) in theoretical and intellectual labor confronts us with urgent ethical and epistemological dilemmas. While their technical capabilities are impressive, their integration into knowledge systems risks reinforcing historical injustices, obscuring questions of authorship, and homogenizing the plurality of critical thought. To engage responsibly with these technologies, a deeper philosophical interrogation is indispensable—one that recognizes the entanglement of epistemology, ethics, and power.

❖ *Bias, Ideology, and the Colonial Archive*

At the foundation of every LLM lies an archive: an aggregation of texts, discourses, and cultural artifacts, collected from the digital remnants of human thought. Yet this archive is not neutral. It is profoundly shaped by the historical violences of race, gender, class, and colonialism (Noble, 2018). The knowledge encoded within LLMs thus carries the sedimentations of these structures, often invisibly.

Training data drawn predominantly from Eurocentric, patriarchal, and capitalist sources privileges certain worldviews while marginalizing others. Even when models are superficially "de-biased" through technical interventions, the deeper ontological bias — the privileging of particular ways of seeing, knowing, and valuing — remains embedded (Birhane, 2021). As Sylvia Wynter (2003) reminds us, the "Man" at the center of modern knowledge systems is a colonial invention, and any uncritical reproduction of his archives risks perpetuating epistemic violence.

Moreover, the algorithmic nature of LLMs compounds this issue. Their goal is to predict the most statistically probable continuation of a given prompt, not to engage in counter-hegemonic critique. In doing so, they tend to reinforce dominant ideologies under the guise of neutrality. The colonial archive is not merely preserved but operationalized, automated, and scaled.

Thus, the ethical question is not merely about "representation" but about the reproduction of epistemic hierarchies. Without conscious intervention, LLMs risk becoming instruments of epistemic coloniality, camouflaging historical exclusions under the sheen of technological progress.

❖ *Intellectual Labor and Authorship*

Another profound concern lies in the reconfiguration of intellectual labor. Traditionally, authorship has implied intentionality, agency, and the existential commitment of a thinker to their ideas. With the involvement of LLMs in generating texts, this relation becomes blurred. What does co-authorship mean when one of the "authors" is a machine without consciousness, desire, or accountability?

Philosophically, authorship is not merely about producing text but about assuming responsibility for its meanings and consequences (Foucault, 1969). It is an act of positioning oneself within a discursive and ethical field. LLMs, lacking any ontological stake in the discourses they produce, cannot fulfill this role. They are, in this sense, radically irresponsible — not through malice, but through structural incapacity.

Moreover, questions of ownership arise. If an LLM generates a novel theoretical framework or a compelling philosophical argument, to whom does it belong? The developer? The user? The model itself? Intellectual property regimes, already entangled with capitalist logics of enclosure and commodification, are ill-equipped to adjudicate these new forms of semi-automated creativity (Crawford, 2021).

This destabilization of authorship also raises deeper questions about the nature of thought itself. Is thinking reducible to textual production? Or does it require, as Hannah Arendt (1958) argued, a situated engagement with the world, a dialogue between inner voice and outer reality, a "two-in-one" that deliberates and judges? If the latter, then LLMs, despite their linguistic virtuosity, remain outside the true domain of thinking — they are, at best, prosthetics for human creativity, not its replacement.

Thus, ethical engagement with LLM-generated theory must involve transparency about human involvement, critical reflection on the limits of machine contributions, and a renewed appreciation for the existential dimensions of intellectual labor.

❖ *Risk of Theoretical Homogenization*

Perhaps the most insidious risk posed by LLMs is the homogenization of critical thought. By their very design, LLMs are oriented toward the reproduction of patterns and probabilities. They are excellent at synthesizing prevailing discourses, but far less adept at producing radical ruptures or genuinely insurgent ideas.

The danger is that LLM-generated theory will tend toward a smooth, palatable simulation of criticality — a "safe" radicalism that gestures toward dissent without threatening fundamental structures of power (Chun, 2021). Decolonial, abolitionist, queer, and insurgent traditions of thought, which often emerge from lived struggle and existential risk, risk being flattened into stylistic tropes, emptied of their world-transformative force.

Theoretical homogenization operates through a subtle double movement: first, by reducing complex and situated traditions to linguistic templates; second, by selecting for outputs that are more likely to be legible, acceptable, and profitable within dominant academic and cultural markets. Over time, this process could lead to a narrowing of the epistemic imagination, where only certain forms of "criticality" — those that are recognizable and marketable — survive.

Moreover, the speed and scale at which LLMs can generate theoretical content exacerbates this problem. In a future saturated with semi-automated philosophy, the slow, arduous, and often painful process of original thinking may be devalued. As Byung-Chul Han (2017) warns, the acceleration of communication often leads to the impoverishment of meaning. The velocity of LLM-generated thought could produce an illusion of abundance while masking an underlying crisis of genuine theoretical innovation.

In this context, preserving the space for radical, situated, and embodied thought becomes not merely an academic preference but an ethical imperative. True criticality demands more than cleverness; it demands confrontation with suffering, with contradiction, and with the possibility of failure.

Toward a Posthuman Theory-Building Practice

As we move further into the age of ubiquitous artificial intelligence, the task of theory-building must itself be reimagined beyond anthropocentric assumptions. A genuinely posthuman theoretical practice demands that we neither romanticize human cognitive exceptionalism nor surrender critical agency to machines. Instead, we must envision new modes of collaboration that are reflexive, pluralistic, and capable of sustaining the disruptive energies essential for emancipatory thought. Toward this end, the engagement with LLMs and other AI systems must be reconceived not as a quest for authoritative answers, but as a situated, critical dialogue that continually interrogates the conditions of its own possibility.

The first ethical shift lies in treating LLMs not as oracles of truth but as provocateurs—machines that, through their re-combinatory capacities, can unsettle habitual modes of thinking and open spaces for critical reflection (Haraway, 1991). In contrast to the naive instrumentalization of AI as a tool for efficiency or discovery, a critical human-AI collaboration would position the machine as an

interlocutor whose utterances demand interpretation, suspicion, and contestation. As Jacques Derrida (1978) reminds us, the very act of reading — of engaging with a text — is an event of meaning-making, saturated with difference and undecidability. Similarly, responses generated by LLMs must be situated within hermeneutic practices that foreground ambiguity, contingency, and alterity.

Such a practice would also demand that we resist the temptation to anthropomorphize machines or to attribute to them agency in any robust sense. Instead, the focus must remain on the socio-technical assemblages within which these systems are embedded — assemblages that reflect and reproduce broader dynamics of power, labor, and knowledge (Suchman, 2007). Theorists, in this model, become not passive consumers of machine outputs, but active interpreters and co-constructors of meaning, constantly aware of the infrastructural and ideological conditions shaping AI's utterances.

Crucially, a posthuman theory-building practice also calls for the reengineering of AI systems themselves. If current LLMs reproduce dominant epistemologies due to the biases inherent in their training corpora, future models must be designed with a reflexive epistemic architecture. This would entail training models not only on canonical texts but on insurgent, marginalized, and pluriversal knowledge systems—what Boaventura de Sousa Santos (2014) calls the "ecologies of knowledges." A genuinely reflexive LLM would not merely generate text according to statistical probabilities; it would be capable of critiquing, questioning, and unsettling its own outputs, offering multiple, sometimes incompatible perspectives rather than a homogenized synthesis.

Of course, the creation of such systems is fraught with challenges. Technical difficulties aside, there remains the deeper philosophical question: Can a machine, lacking existential finitude, embodiment, and lived suffering, truly engage in critique? Some, like Bernard Stiegler (2016), suggest that technics are not external to human thought but constitutive of it, implying that carefully designed AI systems could participate in the ongoing evolution of human-noetic structures. Others caution that without the phenomenological depth of human experience, machine "critique" risks becoming a hollow simulation. Regardless, the project of designing more reflexive AI is less about achieving full machine autonomy and more about enriching human-AI interactions, making them more plural, critical, and unpredictable.

Finally, rethinking theory-building in the posthuman age requires a radical reinvention of theory education itself. The task is not merely to teach students to use AI tools effectively, but to cultivate in them a critical stance toward machines. This involves a pedagogy of suspicion — an insistence that every machine-generated utterance is an occasion for philosophical questioning, not passive acceptance (Freire, 1970). Theory students must learn to engage LLMs dialogically, treating them as neither gods nor slaves but as enigmatic companions whose offerings must be sifted, interrogated, and reimagined.

This new pedagogy must also foreground the ethical and political stakes of human-machine collaboration. Students must be sensitized to the ways in which AI systems can reproduce or challenge structures of oppression, and trained to intervene creatively in the architectures of knowledge production. They must see themselves not merely as users of technology but as agents capable of reshaping the socio-technical imaginaries within which theory unfolds.

In sum, a posthuman theory-building practice is not simply about the technological augmentation of human thought. It is about cultivating new relations to machines, to knowledge, and to ourselves. It demands a humility before the alterity of the nonhuman, a critical vigilance against the seductions of technocratic rationality, and an unwavering commitment to the unfinished project of liberation. In this unfolding landscape, the theorist becomes not a solitary genius, but a node in a complex, dynamic, and contested ecology of thought — an ecology in which machines, humans, histories, and futures are entangled in ways we are only beginning to understand.

Conclusions

Large Language Models (LLMs) are no longer peripheral computational tools; they have evolved into emergent epistemic agents that participate, however asymmetrically, in the production of meaning. Their capacity to recombine, simulate, and innovate upon vast textual traditions forces

a profound rethinking of foundational categories such as author, theorist, and thinker. In their presence, authorship becomes an assemblage, theorization an event of human–nonhuman entanglement.

Yet, this hybridization does not imply the obsolescence of human critical agency; rather, it demands its radical reinvention. Future social theory must grapple not only with the ontological challenges posed by machinic cognition but also with the ethical, political, and epistemological reconfigurations it necessitates. It must learn to navigate a landscape where thought is no longer the exclusive province of the embodied, suffering subject, but a distributed phenomenon emerging across networks of human and algorithmic actors.

True emancipation, however, will not emerge from either uncritical embrace or anxious rejection of machine agency. It will arise through practices of critical integration—through engagements that foreground pluralism, historicity, and reflexivity, ensuring that the theoretical enterprise remains open to insurgent knowledges, dissenting imaginaries, and transformative futures. In this unfolding epoch, to theorize is to think with and against the machine, forging new possibilities for understanding and for liberation in a world increasingly shaped by the posthuman condition.

Future Implications and Research Directions

As the distinction between human and machinic cognition continues to erode, the necessity for a critical, creative, and ethically-informed integration of Large Language Models (LLMs) into social-theoretical frameworks becomes ever more pressing. The future trajectory of this technological evolution must not simply respond to change but actively shape it toward emancipatory and transformative goals.

Firstly, there is an urgent need for the development of ethical LLMs explicitly designed to foster critical thought. Unlike commercially-driven models entrenched in neoliberal and colonial data infrastructures, future iterations must be conceived as open-source, decolonial, feminist, and reflexive in their architecture. Such systems should not simply replicate dominant epistemologies but should instead cultivate dissenting imaginaries, counter-hegemonic narratives, and pluralistic forms of social theorization (Birhane, 2021; Noble, 2018). Constructing these ethical models requires interdisciplinary collaboration among philosophers, computer scientists, decolonial scholars, and critical theorists, who must co-create frameworks that challenge prevailing biases while promoting intellectual diversity and social justice.

Secondly, the emergence of hybrid cognition necessitates a radical rethinking of pedagogy. Traditional forms of critical theory, with their emphasis on hermeneutics and historical consciousness, must now be supplemented with proficiency in machine reasoning, algorithmic critique, and networked epistemologies. Posthuman pedagogy, in this context, should aim not merely to teach students how to use LLMs but to cultivate a form of relational intelligence—one that enables them to think with, against, and beyond the machine. This approach would foster an educational framework that is simultaneously critical, creative, and ethically grounded, preparing future scholars to engage with hybrid intelligences in ways that are intellectually and socially responsible (Hayles, 1999).

Thirdly, the public sphere—the traditional site of rational-critical debate—is poised for profound transformation. AI-mediated forums, decentralized knowledge platforms, and machine-augmented discourses present unprecedented opportunities for the democratization of knowledge and the emergence of new counter-publics. However, achieving a truly post-anthropocentric model of enlightenment requires vigilance against the homogenizing tendencies of algorithmic governance and the potential capture of discourse by hegemonic forces. Future research must investigate how AI can serve as a tool for expanding democratic participation in knowledge production, while simultaneously guarding against the reproduction of exclusionary practices and structural violences inherent in earlier forms of intellectual and political power.

In this emergent landscape, social theory stands at a critical crossroads: either it will retreat into nostalgic humanism, or it will boldly reimagine its purpose in collaboration with emerging epistemic actors. The challenge here is not merely technological, but deeply existential. It calls for a renewed commitment to critical thought, ethical responsibility, and radical imagination as we navigate the

complexities of hybrid intelligences and reconfigure our collective engagement with knowledge and society.

References

- Adorno, T. W. (1973). *Negative dialectics* (E. B. Ashton, Trans.). Routledge.
- Adorno, T., & Horkheimer, M. (2002). *Dialectic of enlightenment* (J. Cumming, Trans.). Stanford University Press. (Original work published 1947)
- Arendt, H. (1958). *The Human Condition*: University of Chicago Press.
- Baudrillard, J. (1981). *Simulacra and simulation* (S. F. Glaser, Trans.). Semiotext(e).
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1-18. <https://doi.org/10.1145/3411764.3445746>
- Benjamin, W. (2019). The Work of Art in the Age of Mechanical Reproduction. In H. Zohn (Ed.), *Illuminations: Essays and Reflections* (pp. 217-252). Harcourt. (Original work published 1935)
- Birhane, A. (2021). Artificial intelligence and the colonial archive. *Journal of Ethics and Information Technology*, 23(2), 123-145. <https://doi.org/10.1007/s10676-021-09592-2>
- Birhane, A. (2021). Decolonial AI: Decoloniality, AI, and epistemic justice. *AI & Society*, 36(4), 1053-1062. <https://doi.org/10.1007/s00146-020-01046-x>
- Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.
- Bourdieu, P. (1977). *Outline of a theory of practice* (R. Nice, Trans.). Cambridge University Press.
- Bourdieu, P. (1990). *The logic of practice* (R. Nice, Trans.). Stanford University Press.
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. W. W. Norton & Company.
- Brennen, B. (2018). *Qualitative Research Methods for Media Studies*. New York: Routledge.
- Bloch, E. (1959). *The Principle of Hope*. (N. Plaice, S. Plaice & P. Knight, Trans.). Cambridge, MA: MIT Press.
- Chalmers, D. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy*. W. W. Norton & Company.
- Chun, W. H. K. (2021). The trouble with theory: Artificial intelligence and the challenge of radical thought. *Theory, Culture & Society*, 38(3), 35-59. <https://doi.org/10.1177/0263276420978303>
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Crawford, K. (2021). *Atlas of AI: Mapping the politics and power of artificial intelligence*. Yale University Press.
- Derrida, J. (1978). *Writing and difference*. University of Chicago Press.
- Durkheim, E. (1895). *The Rules of Sociological Method* (S. A. Solovay, Trans.). Free Press.
- Epstein, J. M. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press.
- Fisher, M. (2009). *Capitalist realism: Is there no alternative?* Zero Books.
- Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press.
- Floridi, L. (2016). *The Ethics of Information*. Oxford University Press.
- Floridi, L., & Chiriatti, M. (2020). Ethics of artificial intelligence: A critique of the value alignment approach. *Journal of the Philosophy of Information*, 15(1), 1-16. <https://doi.org/10.1007/s11023-019-09451-x>
- Fraser, N. (2008). *Scales of justice: Reimagining political space in a globalizing world*. Columbia University Press.
- Fraser, N. (2014). *Fortunes of feminism: From state-managed capitalism to neoliberal crisis*. Verso.
- Freire, P. (1970). *Pedagogy of the oppressed*. Continuum.
- Foucault, M. (1969). What is an author? In *The archaeology of knowledge* (pp. 159-167). Pantheon Books.
- Gadamer, H.-G. (1975). *Truth and method* (2nd ed.). Continuum.
- Gebru, T., Gebru, S., & Mitchell, M. (2021). Data, discrimination, and debiasing: A study of bias in machine learning. *Journal of Ethics and Technology*, 7(2), 1-13.

33. Haraway, D. (1991). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, cyborgs, and women: The reinvention of nature* (pp. 149-181). Routledge.
34. Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
35. Han, B. C. (2017). *The burn-out society*. Stanford University Press.
36. Heidegger, M. (1927). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Blackwell Publishing, 1962.
37. Kitchin, R. (2014). Big data and human geography: Opportunities, challenges, and risks. *Dialogues in Human Geography*, 4(3), 1-12.
38. Lazer, D. M., Pentland, A., Adamic, L. A., Aral, S., Barabási, A. L., Brewer, D., ... & Christakis, N. A. (2009). Computational social science. *Science*, 323(5915), 721-723. <https://doi.org/10.1126/science.1167742>
39. Marcuse, H. (1964). *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Beacon Press.
40. Merleau-Ponty, M. (1945). *Phenomenology of perception* (C. Smith, Trans.). Routledge, 2012.
41. Marx, Karl. (1867/1990). *Capital: A Critique of Political Economy, Volume I*. (Ben Fowkes, Trans.). London: Penguin Classics.
42. Nietzsche, F. (1887). *On the genealogy of morals* (W. Kaufmann, Trans.). Vintage, 1967.
43. Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
44. Ricoeur, P. (1984). *Time and narrative* (K. McLaughlin & D. Pellauer, Trans.). University of Chicago Press.
45. Santos, B. de S. (2014). *Epistemologies of the South: Justice against epistemicide*. Paradigm Publishers.
46. Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424.
47. Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press.
48. Stiegler, B. (2010). *Taking care of youth and the generations* (D. Ross, Trans.). Stanford University Press.
49. Stiegler, B. (2016). *Automatic society, volume 1: The future of work*. Polity Press.
50. Wynter, S. (2003). Unsettling the coloniality of being/Power/Truth/Freedom: Towards the human, after man, its overrepresentation—An argument. *The New Centennial Review*, 3(3), 257-337. <https://doi.org/10.1353/ncr.2003.0022>
51. Weber, Max. (1922/1978). *Economy and Society: An Outline of Interpretive Sociology*. (G. Roth & C. Wittich, Eds.). Berkeley: University of California Press.
52. Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.