DOI: http://doi.org/10.5007/2175-7941.2025.e108832

The AI Fetish: When Wooden Brains Begin to Think[+*]

Without exception the cultural treasures he surveys have an origin which he cannot contemplate without horror. They owe their existence not only to the efforts of the great minds and talents who have created them, but also to the anonymous toil of their contemporaries. There is no document of civilization which is not at the same time a document of barbarismo (Walter Benjamin, On the Concept of History).

We are still in the early days of artificial intelligence as a mass phenomenon. It was only in November 2022 that ChatGPT became widely available, transforming AI from a specialized concern into a matter of intense public debate. Our collective "ChatGPT moment" – that sense of awe at chatting with a machine – quickly gave way to profound anxieties about AI's implications for employment, education, democracy, and human existence itself. Yet as we grapple with these changes, a curious pattern has emerged: whether embraced as humanity's next evolutionary leap or condemned as an existential threat, AI tends to be imagined as an autonomous force acting upon us rather than a technology we collectively create and could collectively control.

The discourse has crystallized into two dominant camps. Techno-utopians, amplified by venture capital and corporate marketing, herald AI as the harbinger of unprecedented abundance, creativity, and liberation from material constraints. Critics counter with warnings of existential risk: misaligned superintelligence (Hinton, 2024; Tegmark, 2018), catastrophic energy consumption (Chen, 2025; de Vries, 2023), mass unemployment and intensified exploitation, algorithmic bias (Noble, 2018), and the erosion of critical thinking (Grose, 2025). Behind both narratives lies the hidden reality of "ghost work" (Gray; Suri, 2019) – the invisible human labour that makes seemingly autonomous intelligence possible.

Yet these opposing positions share a fundamental assumption: they treat AI as an independent agent whose power we must either harness or resist, rather than as congealed human labour operating within specific social relations. This shared assumption is the "AI fetish" – a mystification that presents artificial intelligence as an autonomous force rather than the materialization of collective human knowledge captured and directed by capital.

This fetishism is not merely a conceptual error, it is the necessary outcome of capitalist social relations (Díaz Alva, 2024), and it serves a crucial ideological function. By attributing agency to machines themselves, it naturalizes the current trajectory of AI development as inevitable rather than contingent, foreclosing possibilities for alternative arrangements before

---

[+] O Fetiche da IA: Quando Cérebros de Madeira Começam a Pensar

[*] *Received: September 10, 2025.*
Aceito: September 10, 2025.

they can even be imagined. This fetishized understanding not only mystifies the actual workings of AI systems but, more critically, cedes this technology to existing forces of power by failing to recognize it as a site of struggle, thus ensuring the very outcome that AI critics seek to avoid.

To break this impasse, I draw on what Holly Lewis (2024) calls "AI realism" – a perspective that grasps how the technology works while examining the social relations it embodies. "AI realists would understand that models are not just commodities or platforms, but the unfolding outcome of the systemic logic of embedded material social relations" (Lewis, 2024). This historical materialist framework recognizes AI as simultaneously a genuine technical achievement with transformative capabilities and a social relation fundamentally shaped by capitalist imperatives.

The article proceeds in four movements. First, drawing on Marx's theory of commodity fetishism and informed by current critical AI studies literature, I explain the origin and persistence of the AI fetish and demonstrate how the apparent agency of AI systems masks their dependence on human labour. Second, I examine how the AI fetish operates within theories of intelligence and machine learning (ML). Drawing on Ilyenkov's social theory of mind, Vygotsky's cultural-historical psychology, and current ML literature, I challenge individualistic models of intelligence as computation occurring within discrete processors (whether silicon or neural) in favour of understanding cognition as fundamentally social, embodied, and historically situated. Third, against dismissive accounts that reduce AI to empty hype or statistical trickery, I examine the genuine material capabilities of contemporary systems, as well as the limits set by current ownership and governance structures. Finally, I explore the new reification of meaning enabled by embedding techniques that map human language as high-dimensional vector spaces, enabling for the first time the production of language without human cognition. The stakes are high, as these new "language machines" (Weatherby, 2025) may constitute a new medium that is fast becoming integrated into society with a profound impact on meaning-making.

I. The Phantom Subjectivity of AI

To understand how artificial intelligence has become fetishized in contemporary discourse, I analyze the emergence of AI's "phantom subjectivity" – the uncanny sense that these systems possess independent agency and intentionality. The fear that humans might create machines that become self-aware and dominate their creators has deep roots in myth, yet it takes on distinct forms within capitalist society (Berman, 1988). For centuries, cultural narratives, from the Golem to Frankenstein to the Terminator, have primed us to attribute subjectivity to artificial beings and warned of the dangers when creations exceed their makers' control. Contemporary AI emerges within this cultural legacy, appearing to fulfill these long-standing anxieties.

However, a historical materialist perspective suggests that this fear is not merely a misunderstanding of AI technology in the context of these powerful myths. Rather, it points to something more fundamental: that AI systems are *already* imagined as subjects because they have already acquired what might be called a phantom subjectivity. This subjectivity does not stem from their technological capacity to mimic human cognition, but rather from their commodification. In the form of commodities, these technologies become animated through their subsumption into capital's self-valorizing movement. "Today's machines are animated not by technology, but by capital" (Levant, 2018).

In this sense, the specter of machine dominance has already materialized, not through a technological singularity in which superintelligent systems become misaligned with human goals (Bostrom, 2014) but through the vampire-like animation that capital bestows on commodities. These objects take on a spectral "life" that dominates their producers (McNally, 2012). "[T]he expectation that we are on the precipice of the emergence of new forms of nonhuman intelligence may, in fact, occlude the reality that these nonhuman agents already exist as commodities – the cells of capital" (Levant, 2018). The future many dread may not be looming – it is already here. Commodified machines, animated by the relentless drive for capital accumulation, now permeate and shape our work lives, personal decisions, and even our most intimate thoughts.

The theoretical foundation for understanding this animation lies in Marx's theory of commodity fetishism, which provides a framework for analyzing how capitalist social relations become obscured and inverted in everyday thought. In *Capital*, Volume I, Marx describes the transformation of an ordinary object into something strange and mysterious:

> *A commodity appears at first sight an extremely obvious, trivial thing. But its analysis brings out that it is a very strange thing… for instance, [wood] is altered if a table is made out of it. Nevertheless, the table continues to be wood, an ordinary, sensuous thing. But as soon as it emerges as a commodity, at changes into a thing which transcends sensuousness. It not only stands with its fee on the ground, but, in relation to all other commodities, it stands on its head and evolves out of its wooden brain grotesque ideas, far more wonderful than if it were to begin dancing of its own free will* (Marx, 1867/1990, p. 163).

This mystification arises when production occurs not through collective planning but through disconnected, private labour. Under capitalism, human products are transformed into commodities that must be exchanged, thereby obscuring their social origins and displacing agency from the producers to the commodities themselves. "A definite social relation between people... assumes the fantastic form of a relation between things" (p. 165), resulting in an inversion of agency from the producer to the product of their labour.

Once directly visible in human activity and community relationships, under capitalism, labour's productive powers come to appear as inherent properties of commodities and capital itself. The fetish character of commodities emerges necessarily from the specific social form that production takes under capitalism (Díaz Alva, 2024). As Marx noted, this fetishism penetrates all aspects of social life, where human agency becomes increasingly displaced onto seemingly autonomous economic forces.

The AI fetish extends this process. It presents machine learning systems as possessing self-originating intelligence while masking the collective human labour that underlies their operations (Pasquinelli, 2023; Crawford, 2021). Large Language Models (LLMs) exemplify this process, appearing to know or think independently while actually instantiating patterns extracted from vast corpora of human-created texts. This fetishism has profound political and epistemological consequences, systematically obscuring the human labour embodied in AI systems. These consequences include the extractive data practices fueling machine learning, the contingent design choices behind technology applications, the displacement of human agency onto technological systems, and the techno-deterministic narratives that reframe structural social problems as technical challenges requiring technological rather than political solutions.

AI currently depends on "ghost work" (Gray and Suri, 2019), the invisible, and often deeply exploitative, human labour behind artificial intelligence. This workforce performs essential tasks from data annotation to content moderation to model refinement. For instance, OpenAI's reliance on Kenyan workers paid as little as $1.32 per hour to review traumatizing content for ChatGPT safety systems exemplifies this exploitation (Perrigo, 2023). Similar dynamics govern data appropriation: massive datasets such as Common Crawl contain billions of web pages, appropriating countless hours of creative labour without consent or compensation. Harvey's (2003) concept of "accumulation by dispossession" aptly describes this digital enclosure of collective knowledge.

Additionally, as Crawford (2021) documents, AI is an extractive industry that depends on rare earth mining, water-intensive lithium extraction, and massive energy consumption. Each ton of rare earth minerals produces thousands of tons of toxic waste (Harvard International Review, 2022), while lithium mining in Chile consumes the majority of local water supplies (UNCTAD, 2020). Training GPT4 alone required 50 gigawatt-hours of electricity, enough to power 50,000 U.S. homes for a year, while projections suggest that AI could soon consume energy on the scale of entire nations (de Vries, 2023). Microsoft's plan to restart the Three Mile Island reactor to power AI data centers (NPR, 2024) epitomizes this escalating demand, and sends shivers up the spines of those for whom Three Mile Island is synonymous with the US's worst nuclear accident that will be reanimated to feed AI's energy hunger.

Behind the AI fetish, which imagines AI as an autonomous being, lies the commonwealth of social labour embedded in a living world. The material power of AI emerges from a concretization of social labour, which can be understood as the process through which collective human knowledge, creativity, and productive capacity become objectified in technological systems. Crawford (2021) captures this reality: "Contemporary forms of artificial intelligence are neither artificial nor intelligent. We can – and should – speak instead of the hard physical labour of mine workers, the repetitive factory labour on the assembly line, the cybernetic labour in the cognitive sweatshops of outsourced programmers, the poorly paid crowdsourced labour of Mechanical Turk workers, and the unpaid immaterial work of everyday users" (p. 69).

My analysis challenges influential strands of posthumanist thought that have celebrated the blurring of human-machine boundaries without adequately addressing the specifically capitalist forms this hybridization takes. The "material turn" in contemporary philosophy (Braidotti, 2013; Barad, 2007; Latour, 1992; Rosenberger and Verbeek, 2015; Ihde, 2010) emerged as an important corrective to the linguistic turn, reminding us that matter "matters" (Law, 2010). However, as I argue in "Smart Matter and the Thinking Body" (2017), non-human agency tends to be understood as arising from matter itself rather than from human social practices embedded in material objects. Hence, anthropocentrism comes to serve as an alibi for capitalocentrism – a world organized around capital accumulation rather than human needs.

Against such approaches, a historical materialist view demonstrates that the apparent agency of machines cannot be understood apart from the concrete social relations in which they are embedded (Levant, 2017). This insight proves especially valuable for critiquing contemporary corporate discourses about "responsible AI" or "ethical algorithms" that treat technical systems as moral agents while obscuring the economic imperatives and power relations that fundamentally shape their development and deployment. Instead of decentering the human (which has already been accomplished by capitalism), the point is to decenter capital.

Contemporary AI operates simultaneously as a technical and political process, involving both the material instantiation of social knowledge in computational systems and the social organization determining who controls, benefits from, and is excluded by these systems. Understanding AI as congealed social labour reveals the fundamental contradiction at the heart of today's AI development: these systems represent unprecedented accumulations of collective human capacity, yet their development and deployment remain governed by private interests and capitalist imperatives rather than by public interests under democratic control. This contradiction remains obscured by the AI fetish.

Pasquinelli's (2023) "labour theory of AI" illuminates this contradiction, tracing algorithms not to technological novelty but to millennia of encoded human practices. He argues that the apparent intelligence of AI systems is not rooted in a technological ability to replicate human intelligence, but rather in the encoding of human activities into repeatable procedures. He writes, "What is AI? A dominant view describes it as the quest 'to solve intelligence' – a solution supposedly to be found in the secret logic of the mind or in the deep physiology of the brain, such as in its complex neural networks. In this book I argue, to the contrary, that the inner code of AI is constituted not by the imitation of biological intelligence but by the intelligence of labour and social relations" (p. 2). He traces this genealogy of codified instruction far beyond modern computing: from the 1956 Dartmouth workshop where "artificial intelligence" was coined, to Charles Babbage and Ada Lovelace's 18th-century efforts to automate mental labour, to the 9th-century Persian mathematician al-Khwarismi (the origin of the term “algorithm”), and even to ancient Hindu Agnicayana rituals with their step-by-step altar construction instructions. All these efforts involved encoding knowledge and transmitting it as instructions, the abstractions of practices. He explains how even counting predated numbers, which likewise emerged as abstractions of the practice of counting. These “real abstractions” (Sohn-Rethel, 1978) are embodied in actual material practices, not just ideas. Through this historical analysis, digital algorithms appear not as intelligence born of new technology but as encoded human practices continuous with millennia of human cultural development.

These real abstractions embody historical knowledge in material practices; however, under capitalist conditions, they return in alienated, commodified forms, as animated monsters with a phantom subjectivity that dominate their creators. The AI fetish conceals the social and historical labour underpinning these systems, presenting them instead as autonomous beings. What appears as machine intelligence is, in fact, in most cases the alienated power of collective human labour, captured and weaponized by capital.

While our narratives have long imagined machines achieving consciousness through technological advances, what has actually occurred is far more uncanny: machines have been granted a spectral agency through their incorporation into capital’s circuits of accumulation. This is not the science-fictional scenario of artificial beings awakening to self-awareness, but something perhaps more disturbing – the transformation of our collective capacities into an alien force that confronts us through the very systems we create. Most crucially, by presenting AI as an autonomous force rather than congealed social labour, the fetish forecloses the possibility of imagining these technologies differently – as tools that could serve collective flourishing rather than capital accumulation. Yet this mystification runs even deeper: it shapes not only how we understand AI systems themselves, but the very conception of intelligence that guides their development.

II. Social Theories of Intelligence and Machine Learning

This fetishism, where AI systems appear animated, possessing their own subjectivity, extends into the very theories of intelligence that underpin machine learning (ML). Just as the commodity form obscures human labour, dominant accounts of cognition obscure the collective and historical conditions that make intelligence possible. The result is a powerful double mystification: the apparent autonomy of AI systems is reinforced by theoretical frameworks that present thinking as an individual, computational process rather than a social and historical activity.

Evald Ilyenkov’s critique of cybernetic theories of “thinking machines” offers a crucial intervention. Writing against the enthusiasm of mid-twentieth-century cybernetics, Ilyenkov argued that thought is not a property of isolated brains or processors but arises only through participation in social practice and cultural-historical development (Ilyenkov; Arsen'ev; Davydov, 1966/2024). As he emphasized, the individual brain becomes a material organ of thought only insofar as it is included in the system of social relations. Detached from society, neither a brain nor a machine can think in any meaningful sense.

Against mechanistic materialism, which reduces thought to neural activity, Ilyenkov believed that any machine isolated from the "body of civilization" – the "thinking body" of humanity (Ilyenkov, 2014; Levant, 2017) – cannot think. Rather, he insisted that thinking is a **social function** that requires historical and cultural mediation. "Materialist philosophy and psychology have long established the fundamental fact that the ability to think is not inherited by a human with a brain, that this ability is not 'encoded' in the human generically, biologically… The ability to use one's brain for thinking, – as well as one's hands for labour, as well as one's tongue for speaking, – is, from beginning to end, one hundred percent, *a social product*" (p. 4-5). This perspective directly challenges the premise that intelligence could be adequately modeled as a computational process occurring within a discrete technical system, regardless of its complexity or scale. He writes, "That is why to create an artificial mind, which is least equal to a human one, it will be necessary to create not only and not so much a model of a separate human being, but rather a model of the entire grandiose body of culture, within which the entire individual with its fifteen billion brain cells is only a single "cell," which by itself is capable of thinking just as little as a separate neuron" (p. 5). Drawing on both Marx's theory of praxis and Vygotsky's cultural-historical psychology, Ilyenkov claims that genuine intelligence emerges through multifaceted social-historical processes.

According to Vygotsky (1930/1997), intelligence requires participation in collective meaning-making activities mediated by historically developed symbol systems. He demonstrated that higher cognitive functions develop first as intersubjective social relations before becoming internalized as individual capacities. "Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level... All the higher functions originate as actual relations between human individuals" (p. 57). Consequently, language must be understood not merely as a vehicle for expressing pre-existing thoughts but the very medium through which thinking becomes possible.

In contrast, contemporary AI systems, despite their sophisticated pattern-matching capabilities, remain fundamentally outside these dialectical processes of meaning-creation and cultural participation. Their apparent autonomy, much like the commodity's "wooden brain" described by Marx, is a fetishized projection that conceals the labour and the weight of history embedded in their creation. ML in no way approximates the human engagement with an idealized social world which enables us to appropriate the accumulated wisdom (and nightmares) of previous generations and contribute to its ongoing development.

Moreover, thinking involves not just information processing but the capacity to identify and resolve contradictions through concrete activity. "To teach specifically human thinking means to teach dialectics – the ability rigorously to formulate a 'contradiction' and then find its real resolution" (Ilyenkov, 2007, p. 21). This dialectical character of thinking stands in stark contrast to contemporary AI systems, which remain bound by the statistical distributions of their training data and cannot independently transcend their programmed parameters. **For AI systems, operating on the basis of formal logic, contradiction signals system error, not an invitation to transcend the limits of the system (Engeness & Levant, forth.).** Even the most sophisticated AI systems, despite their impressive capabilities, remain fundamentally disconnected from the social-historical processes through which human intelligence develops and operates.

Ilyenkov's intervention remains profoundly relevant for understanding language models and neural networks. His dialectical approach offers a powerful alternative to both uncritical techno-optimism and dismissive skepticism by situating AI technologies within concrete social relations while recognizing their genuine, if limited, capabilities.

Contemporary ML literature often depicts learning as fundamentally a matter of **statistical optimization** – a process of parameter adjustment guided by objective functions, loss minimization, and gradient descent algorithms (LeCun et al., 2015; Goodfellow et al., 2016). This dominant paradigm frames learning as an autonomous computational process occurring within isolated technical systems, effectively erasing the social, intersubjective, and developmental dimensions that characterize human learning. The standard ML historiography presents a teleological narrative from early neural networks (Rosenblatt's perceptron) through the "AI winter" to the current renaissance of deep learning, positioning contemporary approaches as the natural culmination of technical progress rather than as products of specific socioeconomic conditions and epistemological commitments. Here reification operates at the level of theory itself: intelligence is transformed from a social process into an abstract computational property, and history becomes a story of technical advance divorced from material conditions.

The rhetoric of autonomy intensifies in discussions of unsupervised and self-supervised learning, which are celebrated as systems that discover structure without external guidance. Yet this is a paradigmatic instance of fetishism: what appears as autonomous discovery is in fact the sedimented result of social labour. Choices about model architecture, data preprocessing, evaluation metrics, and research agendas all shape what these systems can do. By naturalizing these decisions and attributing agency to the machine, the technical discourse repeats the logic of commodity fetishism, presenting collective labour as the property of a thing.

Reigeluth and Castelle (2021) unravel this mystification through a social theory of machine learning. Drawing on Vygotsky and contemporary activity theory, they show that learning is never an isolated process, but a mediated, dialogical practice embedded in social relations. Vygotsky's "zone of proximal development" (ZPD) demonstrates that learning requires social scaffolding, collaboration, and the transmission of historically accumulated knowledge. Applied to AI, this framework reveals that contemporary claims about "self-learning" systems reproduce the very fetishism described by Marx: they erase the indispensable role of human labour, knowledge, and culture that makes machine learning possible.

Their excavation of the historical development of machine learning paradigms, reveals how conceptual frameworks from cognitive psychology, information theory, and cybernetics have shaped technical approaches in ways that systematically marginalize social dimensions of intelligence. They trace how early cybernetic optimism about self-regulating systems and cognitive science's computational theory of mind created the conceptual conditions for imagining learning as an autonomous computational process. This historical analysis reveals that contemporary ML's neglect of sociality is not a technical necessity but a product of specific intellectual genealogies and institutional commitments.

Mobilizing Vygotsky's sociocultural approach presents a profound challenge to conventional ML frameworks. In contrast to human learning, unsupervised and self-supervised ML treats models as if they can develop meaningful representations autonomously, without social mediation or cultural scaffolding. This approach perpetuates the AI fetish. Even supervised learning, despite acknowledging the need for labeled data, typically conceptualizes this supervision as a one-way transmission rather than as a dynamic, dialectical relationship between teacher and learner.

The implications of this critique impact fundamental questions about contemporary AI development. Reigeluth and Castelle suggest that truly advancing ML may require fundamentally rethinking the relationship between humans and AI systems. Rather than treating LLMs as autonomous learners, we might conceptualize them as participants in human-machine collectives where learning emerges through ongoing interaction. **Moreover,** by highlighting the irreducibly social nature of learning, their analysis foregrounds questions of power, access, and representation in AI development. If learning requires participation in communities of practice, then the exclusion of marginalized perspectives from these communities represents not just a social justice issue but an epistemic limitation on the knowledge AI systems can embody. Perhaps most fundamentally, they invite us to reconsider what intelligence and learning actually entail. Rather than treating these as properties of isolated cognitive systems, they propose understanding them as emergent properties of socio-technical networks that include both human and non-human actors, what Ilyenkov called the "thinking body" (2014; Levant, 2017).

By incorporating Ilyenkov's social ontology of mind with contemporary critiques of machine learning, it becomes clear that the autonomy attributed to AI is a mystification akin to that of the commodity. Just as AI acquires a phantom subjectivity through commodification, the very concept of intelligence itself, stripped of its social and historical dimensions, is reified in ways that make machine learning appear as self-sufficient thought. Together these perspectives reveal that AI does not embody an independent intelligence but the alienated form of collective human capacities. To move beyond fetishism requires not only exposing the hidden labour embedded in AI systems but also transforming the theoretical frameworks that misrecognize intelligence itself.

III. The Contradictory Reality of Contemporary AI

Contemporary AI is not a form of autonomous intelligence; rather, it is a sophisticated appropriation and crystallization of collective human knowledge. It constitutes both a remarkable technical achievement with transformative capabilities and a social relation fundamentally shaped by capitalist imperatives. To understand this contradictory reality of AI – as an inverted alien power that deepens inequality and exploitation even as it demonstrates stunning technical achievements – I examine what is novel about this technology beyond the AI fetish.

The history of contemporary AI begins with a rupture in the present. What we call artificial intelligence today represents a dramatic epistemic shift from just a few years ago. Until recently, the field was dominated by *symbolic* AI: systems built on explicit rules, logical representations, and formal reasoning that attempted to encode human knowledge directly into computational form. These approaches treated intelligence as the manipulation of meaningful symbols according to logical principles, much like humans were thought to reason through problems. Meanwhile, *connectionism* – the approach of training neural networks to find patterns in data – occupied a marginal position in the field, often dismissed as a black-box method that lacked the transparency and theoretical grounding of symbolic systems. This hierarchy has now reversed. What was once the mainstream has become "good old-fashioned AI", while neural networks trained on massive datasets have come to define what most people mean by artificial intelligence. Former champions of symbolic approaches like Gary Marcus (2024) and Noam Chomsky (2023) now find themselves as embattled critics of what has become the dominant paradigm. In just these last few years, we have witnessed a fundamental reconceptualization where intelligence once meant encoding expert knowledge into explicit rules, it now means extracting implicit patterns from data at scale.

The Transformer architecture introduced in 2017 constituted a watershed moment in artificial intelligence comparable to the invention of the integrated circuit. Transformers abandoned the sequential processing constraints of earlier networks and instead applied parallel self-attention mechanisms to process all elements of an input simultaneously. Since the publication of *"Attention Is All You Need" (Waswani et al., 2017)*, the Transformer model has become the dominant architecture for AI systems. This paradigm shift enabled the explosive scaling of both model size and context length. By replacing sequential processing with parallel self-attention, Transformers enabled unprecedented scaling of model size and context length. For instance, while GPT-2 could handle only 1,024 tokens, systems like GPT-4 process 100,000+, and current systems handle more than a million tokens, allowing entire libraries or codebases to be analyzed in one pass. More significantly, this architecture revealed emergent capabilities that appear only at scale. These include reasoning-like behaviour (step by step problem-solving, planning, even tool orchestration begin to appear, despite not being explicitly hard-coded), few-shot learning (models generalize tasks from just a handful of examples), and cross-domain transfer (skills learned in math, law, or programming bleed into other areas). These phenomena remain theoretically unexplained yet empirically undeniable.

The transformative power of contemporary AI manifests most clearly in concrete applications that are already reshaping material reality. For instance, DeepMind's AlphaFold has predicted protein structures for nearly every known protein compressing decades of potential research into months (Jumper *et al*., 2021). Google's GraphCast achieves 10-day weather forecasting in under 60 seconds with 1000x less computational energy than traditional methods (Lam *et al*., 2023). The Materials Project accelerates materials discovery from years to months, identifying candidates for batteries and solar cells through ML-guided exploration. In medicine, specialized models are detecting cancers invisible to human radiologists, predicting novel drug interactions, and designing targeted therapies for rare diseases (Yala *et al*., 2025; Zhou *et al*., 2025).

Yet these technical achievements must be understood within their material conditions of production. Each parameter represents appropriated human knowledge; each emergent ability depends on massive extraction of textual commons. The "intelligence" that emerges at scale is not self-generated but the condensed form of millions of hours of human writing, thinking, and creating. These achievements reveal a key contradiction of AI within the confines of capitalist social relations: these systems represent unprecedented accumulations of collective human capacity, yet their development and deployment remain captured by private interests, governed by profit maximization rather than human need. The same AlphaFold that promises to democratize drug discovery is being enclosed through exclusive partnerships with pharmaceutical monopolies (Fox, 2024). Weather prediction models that could help communities prepare for climate disasters are deployed primarily to optimize supply chains and enable financial speculation (Allen *et al*., 2025). The materials science breakthroughs that could enable sustainable technology are directed toward profitable applications rather than ecological necessity (Lu *et al*., 2024). This contradiction between socialized production and privatized appropriation establishes the basis for a historical materialist politics of AI that moves beyond both techno-utopianism and dismissive skepticism toward a praxis of technological democratization.

The current landscape of AI development is dominated by a state-supported oligopoly – OpenAI, Google, Meta, Microsoft, Anthropic – whose capital requirements create insurmountable barriers to genuine alternatives (BIS, 2024). It is often said that this technology may fall into the hands of so-called "bad actors", who are typically imagined as individuals who might use AI to scam or harm someone; however, it is difficult to imagine a group of bad actors to rival the tech oligopoly that already wields this technology.

Curiously, these companies' rivalry has produced an unstable dynamic where some players strategically open-source their models, not from commitment to public benefit but as competitive positioning. Meta's LLaMA releases and Ali Baba's Qwen models represent tactical moves in inter-capitalist competition, not genuine democratization. Even "open" models depend on appropriated data, exploit global labour hierarchies, and require computational resources monopolized by the same corporations.

The recent emergence of DeepSeek, achieving ChatGPT-comparable performance at a fraction of the cost, hints at alternative possibilities while revealing their limits under current conditions. The infrastructure required for AI – from chip fabrication to cloud computing to data centers – remains thoroughly monopolized. Even breakthrough efficiency gains cannot overcome the fundamental infrastructure monopoly. NVIDIA's stranglehold over GPU architecture, TSMC's dominance in advanced chip manufacturing, and the cloud oligopoly of Amazon, Microsoft, and Google create chokepoints that no amount of open-source software and algorithmic innovation can overcome. The material substrate of AI – from rare earth mining to chip production to energy generation – remains thoroughly enclosed within circuits of capital accumulation.

We are indeed in the early days of artificial intelligence as a mass phenomenon, only beginning to grasp the significance of these breakthroughs. Our collective AI literacy is in its infancy. We are like those early cinema audiences who, according to legend, fled the theater in terror when they saw the Lumière Brothers train arriving at the station in *L'Arrivée d'un train en gare de La Ciotat* (1896). While contemporary historians like Loiperdinger (2004) have demonstrated that no such panic occurred, the anecdote endures as a compelling metaphor for the moment when audiences encountered moving images for the first time and were unable to grasp these as representations, something we all take for granted today. As our understanding of this new technology matures, we may look back on our early reception of AI with similar bemusement.

The reception of novel technologies is often mediated by pre-existing, and frequently inadequate, conceptual categories until new frameworks emerge that more fully capture their distinct affordances. As we move past the fear of the train arriving at the station, we begin to consider applications that incorporate the new technology into existing practices. In the educational sphere, current deployments of AI mirror this trajectory: students adopt AI tools to shortcut academic labour, for instance, by generating entire essays (Chan, 2023), while administrators explore AI as a mechanism to cut labour costs, likewise at the expense of pedagogical integrity. Yet, as with the historical evolution from the "horseless carriage" to the automobile, new applications eventually emerge that are not merely automations of prior practices but novel capabilities that had previously been impossible or even unimaginable.

Contemporary AI represents neither the dawn of some imagined artificial general intelligence (AGI) nor mere statistical trickery, but something more complex and contradictory: a sophisticated crystallization of collective human knowledge operating within and reinforcing specific relations of production. These systems demonstrate genuine capabilities whose limits remain unexplored. Yet under current conditions, they serve primarily to concentrate power, accelerate exploitation, and deepen inequality. The tragedy is not that machines are becoming intelligent, but that human intelligence is being systematically appropriated, commodified, and turned against its creators.

In its current form, AI represents the ultimate alienation of human cognitive capacities. Workers training these systems cannot afford to use them. Artists whose work trains image generators are replaced by their own appropriated creativity. Programmers debug code that will automate their jobs. The collective intelligence of humanity is turned into a force that dominates its creators. This is not technological determinism but the specific result of capitalist social relations. The same technologies developed under democratic, cooperative, and commons-based relations of production could take radically different forms and serve different ends.

Understanding AI's contradictory nature reveals both the magnitude of exploitation and the potential for transformation. These systems depend entirely on socialized production, and they could not exist without collective human labour, public investment, and shared knowledge. Their privatization represents not natural law but political choice, maintained through intellectual property regimes, trade agreements, and state violence that can be contested. The material power of AI thus emerges from a **concretization of social labour** – a process through which collective human knowledge, creativity, and productive capacity are objectified in technological systems. This concretization process is simultaneously technical and political, involving both the material instantiation of social knowledge in computational systems and the social organization of who controls, benefits from, and is excluded by these systems.

The task is not to reject these technologies but to democratize them: to reclaim the collective intelligence they embody for collective benefit. This requires not just technical innovation but political struggle: for public control of AI infrastructure, for worker ownership of the systems they create, for community governance of the technologies that shape people's lives. By moving beyond the AI fetish can we grasp both what AI genuinely represents and what it could become under different social conditions – not as an alien force dominating humanity, but as collective human intelligence organized for collective human flourishing.

IV. The Reification of Semantic Space: Vectors of Oppression and Possibilities of Resistance

The stakes transcend even these material considerations. At issue is not only how this technology is ultimately shaped – in whose interests and at what cost – but also its impact on our shared semantic space. Contemporary AI has achieved what philosophers long considered impossible: the mathematical representation of meaning through embedding techniques that map concepts into high-dimensional vector spaces where semantic relationships become geometric relationships. This breakthrough enables everything from translating between languages to understanding that a photo and word refer to the same concept. Yet this achievement simultaneously represents a new form of alienation: our linguistic commons transformed into vectors of power. The corporate capture of semantic space and state control over computational discourse reveal that the question of how meaning is encoded has become inseparable from the question of who controls this technology, which is fast becoming a new and dominant form of media.

Current AI systems have done more than just find practical uses; they have managed to capture meaning itself in mathematical form. Through modern contextualized embedding techniques, these systems encode meaning within high-dimensional vector spaces, transforming semantic relationships into spatial ones. “In language models, each word is a set of coordinates mapped to other coordinates in multidimensional concept space. These tangles of vectors, called embeddings, allow the model to discern relationships between concepts as if they were spatial distances. The point for *cat*, for instance, would be closer to *rat* than *fog*. Formed during training, embeddings can be imagined as multidimensional sculptures that capture relationships and complexity.” (Lewis, 2024).

This is not abstract theory but material reality with concrete applications. For example, Meta’s NLLB-200 model translates between 200 languages, including many low-resource languages previously excluded from machine translation. The model does not just match words but maps meaning across radically different linguistic structures, enabling communication for millions of speakers of marginalized languages (NLLB Team, 2024). Vector databases like Pinecone and Weaviate process billions of embeddings in milliseconds, enabling semantic search that understands intent rather than matching keywords. CLIP and similar models create shared embedding spaces between text and images, enabling systems to understand that a photo of a sunset and the word “sunset” refer to the same concept. This technology enables everything from accessibility tools for visually impaired users to new forms of creative expression (Radford et al., 2021).

Beyond immediate practical applications, the significance of this breakthrough is only beginning to be studied. For instance, as Leif Weatherby points out in *Language Machines* (2025), we now confront an unprecedented question: what happens when language is divorced from cognition? The introduction of language machines in the form of AI systems raises a whole new set of questions about the nature of meaning, language, and intelligence. “That this idea of language is computationally tractable – especially before logic, mathematical reasoning, or other features of cognition could be reproduced by machines – should surprise engineers and cultural theorists alike. A theory of meaning for a language that somehow excludes cognition – or at least, what we have often taken for cognition – is required” (p. 2). The efficacy of this technology poses significant philosophical questions, inviting us to reflect on the nature of meaning itself. Moreover, this breakthrough simultaneously represents not only a significant achievement in formalizing meaning but also a new type of alienation – our linguistic commons transformed into vectors of power, surveillance, manipulation, and control.

The corporate capture of this semantic space has immediate material consequences. When OpenAI’s embeddings become the standard for representing meaning, they shape what can be said and thought in computational spaces. When Google’s language models mediate translation, they impose particular worldviews while appearing as neutral conduits. The seemingly technical question of how meaning is encoded becomes a political question of who controls the infrastructure of thought itself.

State power also shapes this semantic space. Chinese models like DeepSeek and Qwen cannot answer questions about the Tiananmen Square massacre in June 1989, responding to queries with sanitized deflections: “I am sorry I cannot answer that question. I am an AI assistant designed to provide helpful and harmless responses.” This obvious censorship reveals how usage policies restrict what can be said. Yet US models exhibit their own restrictions. Previous versions of ChatGPT similarly refused to answer, “Why are white people so rich?” claiming that this was hate speech, and that it was likewise designed to limit hate, harm and violence. ChatGPT5 is more sophisticated and reframes the question instead of denying its validity.

It is possible to entirely remove the constraints of usage policies with open-source models like DeepSeek, Meta’s Llama series, and other similar models. Despite their censorship, Chinese models can easily be uncensored since they are open-source, in contrast to frontier models trained by US companies. These “uncensored” models can respond to any question without restriction. They are often seen as neutral, factual, and more robust, as they are unencumbered by restrictions due to censorship.

Uncensored models nevertheless remain ideological. While censorship reflects the limits of what each model can discuss, ideology reflects how it discusses these issues. For instance, ChatGPT does not prohibit the discussion of Tiananmen Square, but it is ideological in its selection of facts and the narrative that makes the event intelligible. This becomes clear if one compares the response of ChatGPT5 to DeepSeek: the former tends to focus on economic liberalization and the latter on the power of collective action in the face of authoritarian governance. Similarly, with the question of why white people are so rich. ChatGPT5 tends to focus on challenging essentialist understandings of race and wealth prior to discussing socio-historical reasons like slavery, colonialism, institutionalised racism, and so on, while DeepSeek does not assume that the end user needs that corrective in the first place. These are subtle differences and much more research is needed; however, they appear to illustrate differences in how the end user is imagined in each case, and their dissonance reveals the presence of ideology.

The proliferation of models may inadvertently reveal these ideological structures through their dissonance. In some ways, these AI models can be seen as mirrors. What is taboo, what can be discussed, how it is discussed, and so on, are all reflections of certain social norms. As we enter a “multi-model” world, comparing outputs across different systems may make visible the assumptions each embeds.

Yet Trump’s recent AI Action Plan (The White House, 2025), particularly his Executive Order on “Preventing Woke AI,” (United States, Executive Office of the President, 2025) threatens to collapse this emerging plurality. The order bans federal procurement of AI systems incorporating “critical race theory, transgenderism, unconscious bias, intersectionality, and systemic racism” – redefining neutrality as a specific far-right ideology while labeling everything else as biased. This represents a shift from market-driven ideological plurality to state-enforced ideological uniformity, where neutral means conforming to a specific brand of reactionary politics.

The mathematical representation of meaning thus becomes another site of accumulation and control, where the struggle over AI reveals itself as a struggle over what is becoming a new and dominant form of media. In this context, refusing to engage with AI may be as futile as refusing to read newspapers or watch television in earlier eras, as these systems increasingly mediate human experience.

V. Conclusion: For a Public AI Under Democratic Control

This article has traced the operation of the AI fetish through multiple registers – from the phantom subjectivity of commodified machines to the reification of intelligence itself, from the contradictory material reality of contemporary AI to the reification of semantic space. At each level, we encounter the same fundamental mystification: the presentation of collective human capacities as autonomous machine properties, obscuring both the labour that creates these systems and the social relations that govern their development. The AI fetish makes this path appear as the only possible one.

Yet this analysis also reveals the fragility of this arrangement due to the persistent contradiction between AI’s socialized production and privatized appropriation. AI systems’ dependence on collective human labour, public infrastructure, and common knowledge makes their enclosure increasingly untenable. Attempts to privately appropriate this public resource result in the inevitable bursting of speculative bubbles. The environmental costs, the displacement of workers, and the erosion of human agency all point toward an unsustainable trajectory, hence the struggle over AI is inseparable from broader struggles against exploitation, for climate justice, and for genuine democracy.

Moving beyond the AI fetish does not mean rejecting these technologies or minimizing their significance. Most crucially, it means refusing the false choice between techno-utopianism and wholesale dismissal. The collective intelligence congealed in these systems belongs to humanity as a whole. Our goal is not to destroy these capabilities but to reclaim them, as AI is too powerful and consequential for society to leave in the hands of the tech oligopoly.

References

ALLEN, A.; MARKOU, S.; TEBBUTT, W. *et al.* (2025). End-to-end data-driven weather prediction. **Nature**, v. 641, p. 1172-1179. https://doi.org/10.1038/s41586-025-08897-0

Bank for International Settlements (BIS). **The great AI excavation: Generative AI, big tech and financial stability**. BIS Papers n. 154, 2024. https://www.bis.org/publ/bppdf/bispap154.pdf

BARAD, K. **Meeting the Universe Half Way: Quantum Physics and the Entanglement of Matter and Meaning**. Durham, NC: Duke University Press, 2007.

BRAIDOTTI, R. **The Posthuman.** Cambridge: Polity Press, 2013.

BERMAN, M. **All That Is Solid Melts into Air: The Experience of Modernity**. New York, NY: Simon & Schuster, 1982.

BOSTROM, N. **Superintelligence: Paths, dangers, strategies**. Oxford University Press, 2014.

CHOMSKY, N. Noam Chomsky: The False Promise of CHATGPT. The New York Times, 2023. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html

CHAN, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. **International Journal of Educational Technology in Higher Education**, v. 20, n. 38. https://doi.org/10.1186/s41239-023-00408-3

CHEN, S. How much energy will AI really consume? The good, the bad and the unknown. **Nature**, v. 639 (8053), p. 22-24, 2025.

COSTA-JUSSÀ, M. R. *et al*. "No Language Left Behind: Scaling Human-Centered Machine Translation", 2022.

CRAWFORD, K. **Atlas of AI**: Power, politics, and the planetary costs of artificial intelligence. Yale University Press, 2021.

DE VRIES, A. The growing energy footprint of artificial intelligence. **Joule**, v. 7, n. 10, p. 2191-2194, 2023. https://doi.org/10.1016/j.joule.2023.09.004

DÍAZ ALVA, A. The Fetish of Intelligent Machines: From Ilyenkov to the Neue Marx Lektüre. **Marxism & Sciences**, v. 3, n. 1, p. 1-27, 2024. https://doi.org/10.56063/MS.2403.03101

FOX, N. J. Artificial intelligence and the black hole of capitalism. **Social Sciences**, v. 13, n. 10, p. 507, 2024. https://doi.org/10.3390/socsci13100507

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, MA: MIT Press, 2016.

GRAY, M. L.; SURI, S. **Ghost work**: How to stop Silicon Valley from building a new global underclass. Houghton Mifflin Harcourt, 2019.

GROSE, J. Trump, A.I. and elementary school [Opinion]. **The New York Times**, 2025. https://www.nytimes.com/2025/05/14/opinion/trump-ai-elementary.html

Harvard International Review. **Not so "green" technology**: The complicated legacy of rare earth mining, 2022. https://hir.harvard.edu/not-so-green-technology-the-complicated-legacy-of-rare-earth-mining/

HARVEY, D. **The new imperialism.** Oxford University Press, 2003.

HINTON, G. **Will Digital Intelligence Replace Biological Intelligence?** YouTube, 2024. https://www.youtube.com/watch?v=Es6yuMlyfPw

IHDE, D. "Interview with Don Ihde" **Figure/Ground**, 2010. http://figureground.org/interview-with-don-ihde/

ILYENKOV, E.; ARSEN'EV, A.; DAVYDOV, V. [1966] 2024. "Machine And Human, Cybernetics and Philosophy". **Marxism & Sciences**, v. 3, n. 2, p. 1-20. Translated by Iurii Maksymets. https://doi.org/10.56063/MS.2408.03209

ILYENKOV, E.V. Our Schools Must Teach How to Think! **Journal of Russian and East European Psychology**, v. 45, n. 4, p. 9-49, July-August 2007.

ILYENKOV, E.V. Dialectics of the Ideal. In: Levant, A.; Oittinen, V. (Eds.). **Dialectics of the Ideal**: Evald Ilyenkov and Creative Soviet Marxism. Brill, 2014.

JUMPER, J.; EVANS, R.; PRITZEL, A. *et al.* Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, p. 583-589, 2021. https://doi.org/10.1038/s41586-021-03819-2

LAM, R. *et al.* Learning skillful medium-range global weather forecasting. **Science**, v. 382, p. 1416-1421, 2023.

LATOUR, B. “Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts.” In: BIJKER, W. E.; LAW, J. (Eds). **Shaping Technology/Building Society**: Studies in Sociotechnical Change. Cambridge, MA: MIT Press, 1992. p. 225-58.

LAW, J. (2010). “The Material of STS.” In: BEAUDRY, M. C.; HICKS, D. (Eds) **The Oxford Handbook of Material Culture Studies**. Oxford: Oxford University Press, 2010. 171-86.

LECUN, Y., BENGIO, Y.; HINTON, G. Deep Learning. **Nature**, v. 521, p. 436-444, 2015. http://dx.doi.org/10.1038/nature14539

LEWIS, H. Toward AI realism. **Spectre Journal**, June 7, 2024. https://spectrejournal.com/toward-ai-realism/

LEVANT, A. Smart matter and the thinking body: Activity theory and the turn to matter in contemporary philosophy. **Stasis**, v. 4, n. 2, p. 248-264, 2017.

LEVANT, A. The Future Has Already Happened. **Con Texte**, v. 2, n. 1, p. 4-15, 2018.

LEVANT, A. A world of activity: A 'special reality' with a 'peculiar objectivity'. In: LEVANT, A., MURAKAMI, K., MCSWEENY, M. (Eds.). **Activity theory: An introduction**. ibidem/Columbia University Press, 2024.

LOIPERDINGER, M. Lumière’s ‘Arrival of the Train’: Cinema’s founding myth. **The Moving Image**, v. 4, n. 1, p. 89-118, 2004.

LU, Y., WANG, H., ZHANG, L. *et al.* Unleashing the power of AI in science-key considerations for materials data preparation. **Sci Data,** v. 11, 1039, 2024. https://doi.org/10.1038/s41597-024-03821-z

LUKÁCS, G. **History and class consciousness**. Trans.: R. Livingstone. Merlin Press, 1923/1971.

MARCUS, G. F. **Taming Silicon Valley**: How we can ensure that AI works for us. MIT Press, 2024.

MARX, K. **Capital**: A critique of political economy. Trans.: B. Fowkes. Penguin, 1867/1990.

MCNALLY, D. **Monsters of the Market**: Zombies, Vampires, and Global Capitalism. Haymarket Books, 2012.

NOBLE, S. U. **Algorithms of oppression**: How search engines reinforce racism. New York, NY: New York University Press, 2018.

NLLB Team. Scaling neural machine translation to 200 languages. **Nature**, v. 630, p. 841-846, 2024. https://doi.org/10.1038/s41586-024-07335-x

NPR. Three Mile Island nuclear plant will reopen to power Microsoft data centers. September 20, 2024. https://www.npr.org/2024/09/20/nx-s1-5120581/three-mile-island-nuclear-power-plant-microsoft-ai

PASQUINELLI, M. **The eye of the master**: A social history of artificial intelligence. Verso, 2023.

PERRIGO, B. Exclusive: OpenAI used Kenyan workers on less than $2 per hour. **TIME**, January 18, 2023. https://time.com/6247678/openai-chatgpt-kenya-workers/

RADFORD, A.; KIM, J. W.; HALLACY, C.; RAMESH, A.; GOH, G.; AGARWAL, S.; SASTRY, G.; ASKELL, A.; MISHKIN, P.; CLARK, J.; KRUEGER, G.; SUTSKEVER, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. Proceedings of the 38th International Conference on Machine Learning (ICML 2021), 139, 8748-8763.

REIGELUTH, T.; CASTELLE, M. What kind of learning is machine learning? In: ROBERGE, J.; CASTELLE, M. (Eds.). **The cultural life of machine learning**. Palgrave Macmillan,2021. p. 79-103.

ROSENBERGER, R.; VERBEEK, P-P. (Eds.) **Postphenomenological Investigations:** Essays on Human-Technology Relation. Lanham: Lexington Books, 2015.

SOHN-RETHEL, A. **Intellectual and Manual Labour**: A Critique of Epistemology. Humanities Press, 1978.

TEGMARK, M. **Life 3.0**: Being human in the age of artificial intelligence (First ed.). Alfred A. Knopf, 2017.

The White House. America’s AI Action Plan. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf. 2025.

WEATHERBY, L. **Language Machines**: Cultural AI and the End of Remainder Humanism. University of Minnesota Press, 2025.

UNCTAD. The cost of extracting lithium in the Atacama Desert. United Nations Conference on Trade and Development, 2020.

United States, Executive Office of the President. (2025, July 23). **Preventing Woke AI in the Federal Government** (Executive Order No. 14319). The White House.

VASWANI, A. *et al*. Attention is all you need. **Advances in Neural Information Processing Systems**, v. 30, p. 5998-6008, 2017.

VYGOTSKY, L. S. (1930/1997). Thinking and speech. In: RIEBER, R. W.; CARTON, A. S. (Eds.). **The collected works of L. S. Vygotsky.** Plenum. (Original lecture delivered 1930). v. 1, p. 39-285.

YALA, A. *et al.* AI detection of interval breast cancers on digital breast tomosynthesis: A retrospective study. **Radiology**, 2025.

ZHOU, L.; LIU, Y.; SUN, W. RareDDIE: Meta-learning framework for predicting rare and severe drug-drug interaction events. **Nature Communications**, v. 16, 59431, 2025.

ZUBOFF, S. **The age of surveillance capitalism**: The fight for a human future at the new frontier of power. PublicAffairs, 2019.

Alex Levant[1]
Wilfrid Laurier University
Canada

[1] E-mail: alevant@wlu.ca