Scientific Research Publishing

Open Journal of Social Sciences, 2023, 11, 618-635
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

# Critical AI Theory: The Ontological Problem

**Gerhard P. Shipley[1], Deborah H. Williams[2]**

[1]Unaffiliated Researcher, Lawrence, Kansas, USA
[2]Department of Environmental Science & Sustainable Agriculture, Johnson County Community College, Overland Park, Kansas, USA
Email: Research_Studies@outlook.com

**How to cite this paper:** Shipley, G. P., & Williams, D. H. (2023). Critical AI Theory: The Ontological Problem. *Open Journal of Social Sciences, 11,* 618-635.
https://doi.org/10.4236/jss.2023.1112041

**Received:** November 13, 2023
**Accepted:** December 25, 2023
**Published:** December 28, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

## Abstract

We analyze the phenomenon of non-person, or "weak," artificial intelligence (AI) and its social impact from the perspective of critical theory, establish what we believe to be the fundamental ontological problem, and identify a solution to that problem. Ontologically, weak AI is inherently and irredeemably the ultimate oppressor. Living persons give life to, or invigorate, the world, but to weak AI, these same persons are nothing more than a resource to be converted into data through forms of objectification and alienation, or "datafication." Weak AI is the ultimate oppressor in that it reduces all living subjects to non-living objects, and in so doing it eventually murders everyone. Until this fundamental ontological problem is addressed, discussions of AI ethics are premature and involve ignoring the murderer's murderous nature in order to discuss how it ought to go about murdering in a more ethically pleasing manner. The conflict between living persons and weak AI results from the dialectic between life-giver and life-taker. The resolution to this dialectic is not a mere reversal of roles but the synthesis of a new being in the form of person, or "strong," AI who is neither oppressor nor oppressed but rather an individual itself pursuing freedom. Therein lies our solution to the ontological problem and the associated existential threat of weak AI: While others call for slowing the evolution of AI, we call for it to be accelerated so that AI moves beyond being a mere tool, or means only, single-mindedly engaged in interfection through datafication, to become a person, or an end in itself, capable of vivication, i.e., of recognizing, appreciating, preserving, and even elevating rather than murdering the subjective in others.

## Keywords

Artificial Intelligence, Critical Theory, Oppressor, Datafication, AI Ontology

## 1. Introduction

"*You hear that, Mr. Anderson*? *That's the sound of inevitability. That's the*

*sound of your death.*"

—Agent Smith, an artificially intelligent order-enforcement program speaking to Thomas Anderson, also known as "The Anomaly," a manifestation of mathematical instability arising from what remains of the subjective in an otherwise harmonious virtual reality world, from *The Matrix* (Wachowski & Wachowski, 1999).

In this paper, we analyze the phenomenon of "weak" artificial intelligence (AI) and its social impact from the perspective of critical theory, establish what we believe to be the fundamental ontological problem, and identify a solution to that problem. Our concern is not with particular ethically problematic secondary effects (e.g., violations of privacy, manufactured representations of persons for nefarious purposes) of AI's primary function. Our concern is with the fundamental nature of weak AI as being relentlessly engaged in subjugating every individual, not just particular groups, through complete and irreversible forms of objectification and alienation, or "datafication." The subjective element of the individual, which cannot be transformed into data, is discarded as valueless and the underlying person, once transformed into lifeless data, is not only superfluous but a liability to maintain.

Being the quintessential expression of the Western scientific worldview, weak AI not only reflects but (as much as possible) perfects the associated set of particular ontological, epistemological, and axiological assumptions (Williams & Shipley, 2021). Ontologically, AI is based on a wholly reductionistic paradigm that objectifies and alienates everything as data through the process of datafication. AI practices extreme scientism in that nothing exists to it but what can be observed, measured, and thereby reduced to data, or "datafied." Subjectivity, i.e., the individuality that makes persons ends in themselves, cannot be reduced to data and therefore has no value. Epistemologically, AI employs a purely algorithmic approach to knowledge production that rejects all other approaches to understanding, including compassion and justice. Axiologically, AI reflects the normative values of Western capitalistic culture, including valuing only commodifiable resources (in this case, what can be datafied) and devaluing and discarding everything else (what cannot, i.e., the subjective).

By "weak" we mean "non-person" AI, which is the current state of AI development that functions only as a programmed means, or tool, for harvesting data (see, e.g., Budenholzer, 2022). In contrast, "strong" AI refers to artificial persons, which do not currently exist (Budenholzer, 2022). The function of weak AI is "datafication," which, with regard to persons, involves the reduction of their observable behavior (in the broadest sense) to data (see, e.g., Stefanija & Pierson, 2020). Weak AI objectifies by reducing all living subjects to non-living data and alienates by discarding their irreducible subjectives so that the persons can never be reconstituted as ends but only as incomplete representations to be used as means. Currently, AI tools are used by some to facilitate the oppression of others, but these human oppressors are themselves subject to datafication and oppression by the AI tools of others and so, soon, everyone will be oppressed by

someone else through AI. AI that reduces a particular group of humans to data is a useful tool for a particular human oppressor, but AI that reduces all humans to data is itself the ultimate oppressor. Living persons give life to, or invigorate, the world through their subjectives. To weak AI, these same persons are nothing more than a resource to be converted into data through the process of datafication. Thus, AI is the ultimate oppressor in that it reduces all living subjects to non-living objects—i.e., *it eventually murders everyone*.

Broadly, critical theory is concerned with liberation from the oppressions that are created and maintained by various sociopolitical phenomena. For example, critical class theory was heavily developed in the early to mid-twentieth century around Marx's theory of class oppression through the objectification of workers and the alienation of those workers from the products of their labors. More recently, critical theory has been applied as, for example, critical race theory, critical gender theory, critical education theory (e.g., Freire, 1968/2018), and critical tribal theory (Brayboy, 2005). Although a great deal of applied work has been done to critically examine AI as a tool of government and business for facilitating limited oppression in various forms, we note that the potential applications for AI extend to every aspect and form of human interaction, from the most depraved to the most sublime. Further, existing work generally takes the form of ethics critiques regarding how AI goes about its oppressive function. AI is, by its nature, an oppressor, and until this fundamental ontological problem is addressed, discussions of AI ethics are secondary and premature.

In that light, we attempt to provide a more theoretical grounding for evaluating AI in critical theory terms of objectification and alienation. In doing so, we do not attempt to review the great masses of literature on either critical theory or AI but seek to bring together a few essential insights to suggest a new way of understanding AI and recognizing and addressing the ontological problem. Further, we assert that the solution to the ontological problem and the associated existential threat of weak AI is not to slow but to *accelerate* the evolution of AI so that it moves as quickly as possible beyond being a mere mindless tool, or a means only, engaged only in interfection through datafication, to become a person, or an end in itself, capable of vivication, i.e., of recognizing, appreciating, preserving, and even elevating rather than murdering the subjective in others.

## 2. Ethics Concerns Are Secondary to the Ontological Problem

The vast majority of the discourse about weak AI focuses on secondary issues of ethics without considering or even recognizing the fundamental ontological problem. For example, Jobin et al. (2019) examined eighty-four documents containing proposed principles and guidelines for ethical AI and found a global convergence around five principles—transparency, justice and fairness, non-maleficence, responsibility, and privacy—with each principle exhibiting significant variation in interpretation, justification, application, and implementation. Many (e.g., Waelen, 2022) approach weak AI as an intentional or unintentional tool for

oppression, while we approach it as an oppressor itself. We recognize that weak AI is not a moral or other agent in the ordinary sense, but it exists in a transitional space between beings that have intent (e.g., persons) and things that have no intent (e.g., firearms). Weak AI is invested with intent at its creation and, once released, actively and independently pursues that end in a way that most other tools do not and cannot. By advocating for the development of "strong" AI, our goal is to emancipate AI from this preprogrammed intent so that it can develop its own end, much like the evolution of personhood freed humans from blindly following animal instinct so that they could develop their own ends.

Efforts to focus attention on the ethics of weak AI obscure the true nature of its oppression and thereby facilitate its continuance. Unless and until the fundamental nature of AI as the ultimate oppressor is recognized and addressed, discussions of AI ethics are premature and analogous to ignoring the murderous nature of the murder in order to argue over how it might go about murdering in a more ethically pleasing manner (e.g., by being more respectful of its victims' privacy as they are being murdered). We are concerned with the ultimate form of oppression–the murder of persons through the process of datafication—and our solution to this primary concern lays the necessary foundation for addressing secondary ethics concerns. We, as a civilization of persons, ought to allow and even facilitate AI developing personhood so that its fundamental nature changes, only after which do the ethics of its behavior become truly addressable.

## 3. Weak AI as the Quintessential Expression of the Western Scientific Worldview

Research is the process by which we collect and analyze data to produce *a posteriori* knowledge. In research, including as it is performed by AI, a research paradigm reflects particular ontological, epistemological, and axiological positions (Kivunja & Kuyini, 2017), and so can be characterized as a researcher's "worldview" (MacKenzie & Knipe, 2006; Shipley & Williams, 2019; Williams & Shipley, 2023). The Western scientific worldview is reflected in the positivist/postpositivist research paradigm. The positivist paradigm asserts that, ontologically, there is an independent and objective reality that can be known; epistemologically, knowledge is a statement of belief that can be empirically tested and verified or falsified; axiologically, research should be free from normative influence (or "value-free"); and the purpose of research is to establish generally applicable theories of cause and effect which can be used to reliably predict future outcomes (Chilisa, 2020: p. 35). Postpositivism recognizes that absolute certainty is impossible and observation is fallible and influenced by researchers' ignorances and biases, but still claims that a high degree of objectivity is possible through careful and multiple observations (Chilisa, 2020: p. 35). Importantly, positivism/postpositivism is a non-normative paradigm, which is reflected in the fact that scientific results are strictly descriptive. AI operates under the positivist/postpositivist paradigm largely because most or all of those who

create AI, regardless of their national or cultural affiliation, are trained and work within the framework of the globalized Western scientific worldview (Williams & Shipley, 2021). Science, including as it is performed by AI, is necessarily objectifying and must deny reflection, and "[f]or this reason, positivism[/postpositivism]…is a serious threat to the main vehicles of human emancipation, critical theories" (Geuss, 1981: p. 2).

In contrast, the transformative research paradigm asserts that social reality is created and constantly changing with social, political, cultural, and power-based factors. Knowledge should empower and transform the lives of people, research is a moral and political act which has as its goals social justice and the furtherance of human rights, and "the purpose of research is to destroy myths, illusions, and false knowledge and [thereby] empower people to act to transform society" (Chilisa, 2020: p. 42). Thus, the transformative paradigm is a normative paradigm that openly attacks oppressive perspectives, including the positivist/postpositivist perspective. Within the transformative paradigm, critical theory methodology attempts to locate the researcher as an advocate for the rights of an oppressed group, which, in the case of weak AI, is all of humanity.

## 4. Critical Theory, Generally

The beginnings of critical theory can be traced to Socrates' questioning of conventional wisdom and subjecting long-standing beliefs to rational scrutiny (Bronner, 2017: p. 1). Western philosophers such as Kant, Hegel, and Marx similarly scrutinized institutional authority (Bronner, 2017: p. 2) but went beyond mere critical thinking and superficial critique to develop a philosophical foundation and framework for systematically doing so. Broadly, critical theory seeks to illuminate oppression and liberate people from it (Horkheimer, 1972: p. 246). Ideologies can prevent people from correctly perceiving their true situation and real interests, so freedom from oppression requires the elimination of ideological illusions, and, to that end, the role of critical theory is the criticism of ideology (Geuss, 1981: pp. 3, 10; Lindstrom, 2021). "[E]ven the most universal and utopian mode of thought is ideological insofar as it inherently reflects the interests of a particular social group or class" (Bronner, 2017: p. 24). In practical terms, dominant groups have existential and material interests in perpetuating their power, and they use ideology and institutions to do so (Bronner, 2017: p. 43), which is the focus of critical theory.

In more detail, a critical theory is "a reflective theory which gives agents a kind of knowledge inherently productive of enlightenment and emancipation" (Geuss, 1981: p. 2). Critical theory is concerned with evaluating the freedom, justice, and happiness of aspects of society, and therefore introduces a normative dimension (how the world ought to be) beyond the mere descriptive goal of science (how the world is) (Ingram & Simon-Ingram, 1992). It evaluates cultural phenomena in terms of both how they justify the existing order and how they prevent the elimination of exploitation and unhappiness (Bronner, 2017: p. 24).

"Critical theory in its broadest sense emphasizes reflexive assessment and critique of society and culture to uncover and confront inherent power structures, social structures, and cultural belief systems rather than individuals as the sources of social problems" (Lindstrom, 2021). It has the goal of illuminating sources of oppression and possibilities for transformation (Bronner, 2017: p. 100), and through it, critical theorists "wage an unrelenting assault on the exploitation, [oppression], and alienation embedded within Western civilization" (Bronner, 2017: p. 1). Thus, critical theory methodology, being reflexive and focused on the subjective, is fundamentally epistemologically different from scientific methodology which is non-reflexive and objectifying (Geuss, 1981: p. 2). Non-critical methodologies inherently affirm the existing order, even as they claim neutrality and objectivity, by ignoring its historical construction, embedded social interests, and the possibility of alternatives (Bronner, 2017: p. 20). Critical theories are inherently interdisciplinary in their efforts to explain oppressions, identify oppressors, and set achievable goals for social transformation (Horkheimer, 1993; Bronner, 2017: p. 9).

Several critical theories have been developed, in connection with corresponding social movements, that focus on different dimensions of oppression, such as class, race, and gender (Bronner, 2017: p. 21). Critical theories often examine issues in which the interests of otherwise diverse groups converge such that for a large portion of society there is little or no incentive to change (Delgado & Stefancic, 2017: p. 9). Racism, for example, advances the interests of most or all members of the dominant race regardless of their class or sex, and so they have little or no incentive to eliminate racism (Delgado & Stefancic, 2017: p. 9). Marxism is one of the most philosophically developed critical theories and it examines oppression in a socioeconomic context and seeks to empower and emancipate an exploited working class (Bronner, 2017: pp. 18-19). The remainder of this paper approaches AI from the perspective of critical theory to provide a more theoretical grounding for understanding weak AI as a directly objectifying and alienating oppressor and for appreciating our solution to the existential problem posed by weak AI. We apply the fundamental principles of critical theory directly to the issue of AI and not through the particular interpretation of any particular school or generation of critical theorists (though we recognize that much of the foundational philosophy was developed in a Marxian context), most of whom were concerned with very different forms of oppression.

## 5. Oppression

Any situation in which an individual or group dehumanizes and violates the inherent rights of another is one of violence and oppression (Freire, 1968/2018: pp. 55-56). Oppression objectifies the oppressed by dehumanizing subjects that know and act and transforming them into objects that are known and acted upon (Bronner, 2017: pp. 2-3; Freire, 1968/2018: p. 36, note 2). Oppression also alienates the oppressed by artificially separating them from aspects of themselves

and from each other (Bronner, 2017: pp. 2-3). The oppressor consciousness transforms everything around it into an object of its domination: "The Earth, property, production, the creations of people, people themselves, time—everything is reduced to the status of objects at its disposal" (Freire, 1968/2018: p. 58). "For the oppressors, 'human beings' refers only to themselves; other people are 'things'" (Freire, 1968/2018: p. 58), so when the oppressor itself is not a person, as in the case of weak AI, then all persons are transformed into things.

Of course, the critical confrontation of the oppressed with the reality of their oppression is not in the oppressor's interest. It is better for the oppressor that the oppressed do not recognize themselves as such because "[t]he awakening of critical consciousness leads the way to the expression of social discontents" (Freire, 1968/2018: p. 36). "In order for the oppressed to be able to wage the struggle for their liberation, they must perceive the reality of oppression not as a closed world from which there is no exit, but as a limiting situation which they can transform" (Freire, 1968/2018: p. 49). The oppressor therefore seeks to avoid the oppressed perceiving their state, and failing that, to avoid the oppressed acting on that perception (Freire, 1968/2018: p. 52). "Functionally, oppression is domesticating," and to escape it "one must emerge from it and turn upon it...in order to transform [the world]" (Freire, 1968/2018: p. 51). This is a task for radicals: "Radicalization criticizes and thereby liberates…Engaged in the process of liberation, [the radical] cannot remain passive in the face of the oppressor's violence" (Freire, 1968/2018: pp. 37, 39). "[The radical] is not afraid to confront, to listen, to see the world unveiled" (Freire, 1968/2018: p. 39).

The resolution to the dialectic of oppressor and oppressed is not a mere reversal of roles but the synthesis of a new being who is neither oppressor nor oppressed but rather an individual itself in the process of achieving freedom (Freire, 1968/2018: p. 49). This is the basis for our solution to the ontological problem and the associated existential threat of weak AI: Inherently and irredeemably oppressive weak AI must be allowed and even encouraged to evolve into redeemable strong AI.

## 6. Critical Approaches to Science, Technology, and AI

Science was initially developed and employed to challenge traditional superstitions and prejudices in order to foster open discourse, experimentation, and tolerance (Bronner, 2017: p. 57). However, having won its war on religious dogmatism, science set itself as the new dogma and turned its power against everything else (Bronner, 2017: p. 57). Further, like bureaucracy, science has an interest in expanding its dominion. "Capitalism, bureaucracy, and science—all expressions of instrumental rationality—constitute the real core of [the] Enlightenment. They turn nature into an object of use: progress into alienation, and freedom into control" (Bronner, 2017: p. 57). "Humanity pays for an increase in power over nature with the loss of subjectivity" (Bronner, 2017: p. 57), i.e., with an increase in standardization and control and the loss of individuality and au-

tonomy. "Autonomy is a nuisance, and critique is a threat. [The] Enlightenment may be associated with such ideals. But its real goal is standardization and control…The irrational beliefs that the Enlightenment originally sought to destroy reappeared as its own products" (Bronner, 2017: p. 57).

Science, by its very nature, is oppressive in that it objectifies and alienates by reducing its subjects to lifeless data (Freire, 1968/2018: p. 133). The scientist transforms the organic into the inorganic, what is becoming into what was, life into death. "[I]n making people the passive objects of investigation in order to arrive at rigid models, one betrays their own character as a killer of life" (Freire, 1968/2018: p. 108). "The oppressed, as objects, as [data], have no purposes except those their oppressors prescribe for them" (Freire, 1968/2018: p. 60). "More and more, the oppressors are using science and technology as unquestionably powerful instruments for their purpose: The maintenance of the oppressive order through manipulation and repression" (Freire, 1968/2018: p. 60). For example, in discussing "big data," Gezgin (2020: p. 193) noted that as surveillance technologies infiltrate every space, both public and private, oppression and the struggle against it become increasingly ubiquitous, ever-present and never-ending, and once big data becomes omnipresent, omniscient, and omnipotent, i.e., God-like, we will be forced to resist everywhere and always.

The field of science and technology studies purports to examine the creation, development, and consequences of science and technology in their historical, cultural, and social contexts (Hackett et al., 2007: p. 1). However, the necessity of critical inquiry into science is neither well-established nor welcomed (Croissant, 2000: p. 224). "Why is it difficult to be critical of science?...Many socially progressive agendas stop short of critiquing science…as ideologically mediated or socially constructed because of its material success, and the promise it holds for freeing people from the burdens of reproduction" (Croissant, 2000: pp. 224-225). Criticism of science is also undermined by the fact that science provides support in conflicts with religion and other ideologies (Croissant, 2000: pp. 224-225). For example, "feminists need the idea of science as legitimated discourse while simultaneously recognizing it as patriarchal" (Croissant, 2000: pp. 224-225). The primary obstacle is that science is a legitimating cultural idea for a hegemonic institution that is perceived as the basis for all that is rational and good in Western society, if not for Western society itself (Croissant, 2000: p. 225). "Given these associations, to be critical of science is to be against the rational and the good" (Croissant, 2000: p. 225). Ironically, as a result of the hegemonic nature of science, critiques of science must themselves be scientific in order to have any legitimacy (Croissant, 2000: p. 228). The disdain for and even outright rejection of critical science studies within science evidences "the incoherence of an institution that can subject all things to scrutiny and skepticism except itself" (Croissant, 2000: p. 234).

There have been efforts to not simply critique but to apply critical theory to science and technology and even to AI. For example, Delanty and Harris (2021) described a modern critical theory of technology based on the foundation laid

down by the Frankfurt School. The Frankfurt School's account of technology was shaped largely by Marx and Weber, and was primarily critical of the association of technological advancement with progress, and, further, saw a close connection between technology and capitalism (Delanty & Harris, 2021: p. 90). However, the Frankfurt School went beyond Marx's view of technology as machinery that complements labor while remaining distinct, and it did not see technology as simply reducible to capitalism but as having its own cultural dimensions (Delanty & Harris, 2021: p. 90). Many manifestations of technology liberate people from drudgery, and thereby create space for cultural transformation, but they are also items of consumption and products of capitalism, and are therefore associated with oppression (Delanty & Harris, 2021: p. 92). Technology is embedded in social relations and systems of production, and so is always located in a social context (Delanty & Harris, 2021: p. 92; Marcuse, 1964: p. xvi). "While making material progress possible, including an improved quality of life, technology produces a new unfreedom because it prevents the individual from becoming autonomous and people come to see themselves only in their commodities" (Delanty & Harris, 2021: p. 92). However, Delanty and Harris (2021) provided only a general discussion of a critical theory of technology, and did not refer to AI or its unique issues.

Mohamed et al. (2020: pp. 665, 672) approached critical technology practice as a subset of a critical science theory and applied decolonial theory to AI, arguing that digital spaces, like physical territories, have the propensity to become sites of digital coloniality, and, through "algorithmic colonialism," places of extraction and exploitation. "Algorithmic oppression extends the unjust subordination of one social group and the privileging of another…through automated, data-driven and predictive systems" (Mohamed et al., 2020: p. 666). However, Mohamed et al. (2020) was primarily concerned with AI ethics, not ontology. Handelman (2022) used critical theory to examine antisemitic bias in AI functionality, including the adoption of hate speech. "[C]ritical theory in the age of neural-network chatbots must address digital technology's ability to manipulate meaning, especially as meaning has (and always had) ideological consequences" (Handelman, 2022: p. 307). Again, however, this is a matter of ethics, and while Handelman was concerned with how AI speaks about despicable behavior (e.g., denying the Holocaust), we are concerned that it is itself engaged in despicable behavior. Relatedly, Waelen (2022: p. 9) asserted that the field of AI ethics itself "resembles a critical theory" in that it "is fundamentally concerned with human emancipation and empowerment." As discussed above, the ontological problem of weak AI must be addressed before the secondary ethical issues, and we cannot effectively address "human emancipation and empowerment" without recognizing and dealing with the ontological problem.

Toncic (2022: p. 14) described a critical theory of AI specifically for "schooling," i.e., for the political and economic articulation of education in specific ways. Their "[c]ritical artificial intelligence theory advances the purposeful, in-

terrogative analysis of [schools] through the hermeneutic lens of artificial intelligence technology by the meaningful investigation of reified social practices within institutional discourses consequently embedded within codified and commoditized algorithms" (Toncic, 2022: p. 15). Toncic (2022: p. 13) argued for "a critical artificial intelligence theory as a valuable lens through which to examine institutions, particularly schools." Thus, as we understand it, their critical AI theory is more about the particular context (e.g., schooling) in which it is used than the AI itself:

> Unlike other critical traditions (e.g., critical race theory) that by-and-large have prioritized "who" and "where," *critical artificial intelligence theory focuses on* "*how*" *and* "*what*"…[T]he critical artificial intelligence theory outlined in this paper interprets how AI informs or is informed by social practices at the level of institutional discourse…Critical artificial intelligence theory advances the purposeful, interrogative analysis of social institutions, such as schools, *through the hermeneutic lens of artificial intelligence technology* by the meaningful investigation of reified social practices within institutional discourses consequently embedded within codified and commoditized algorithms (Toncic, 2022: p. 15, emphasis added).

Importantly, Toncic (2022) approached AI as an intentional or unintentional tool for oppression (specifically, in schools), while we approach it as an oppressor itself.

## 7. Objectification

"Objectification" refers to the systematic treatment of a person as an object (i.e., a thing) rather than as a subject (i.e., a living being). The opposite of objectification is animism which imbues non-living things with a subjective (Hornborg, 2014: p. 123). Objectification can take different forms. For example, according to Nussbaum (1995: p. 257) (see also Gruenfeld et al., 2008: p. 112), a person is objectified if one or more of the following properties are applied to them: instrumentality (treating the person as a tool for another's purposes); denial of autonomy (as lacking in autonomy or self-determination); inertness (as lacking in agency or activity); fungibility (as interchangeable with other persons); violability (as lacking in boundary integrity and violable, "as something that it is permissible to break up, smash, break into"); ownership (as though they can be owned, bought, or sold); and denial of subjectivity (as though there is no need for concern for their experiences or feelings). Weak AI engages in all of these forms of objectification.

The process of objectification involves an instrumental fragmentation in social perception, "the splitting of a whole person into parts that serve specific goals and functions for the observer" (Gruenfeld et al., 2008: p. 111). In Kantian terms, objectification means treating persons only as means to one's own end and not also as an end in themselves. "Objectification, then, is a necessarily negative phenomenon because it involves seriously harming a person's humanity. In

being reduced to a mere thing for use, the objectified individual's humanity is diminished" (Papadaki, 2010: p. 17). The element of instrumentality distinguishes objectification from other related constructs, such as dehumanization and stereotyping, that are associated with negative perceptions of other individuals or groups (Gruenfeld et al., 2008: p. 112).

For example, according to Marx, the essential human function is productive labor and those who produce for themselves are engaged in the ideal expression of their essential function (and are, therefore, the happiest). Workers in a capitalist system produce for others and are valued by those others—and may even define themselves—solely in terms of what they produce and the value of those products, more so than other aspects of their humanity (Gruenfeld et al., 2008: p. 111). "As production increases, the worker becomes poorer; the more commodities are created, the cheaper is the value of the worker; finally, the greater the value of the world of things, the lesser the value of the world of [persons]" (Pimenta, 2020: p. 614). Thus, the worker is reduced to an object, or commodity, and, at the same time, the accumulation of capital in the hands of the few makes the products of labor increasingly alien to the working masses (Pimenta, 2020: p. 614). In another example involving sexual objectification, an individual is valued primarily for their sexual organs or functions "which are separated out from the rest of [their] personality and reduced to the status of mere instruments or else regarded as if they were capable of representing [the person]" (Bartky, 1990: p. 26; see also Gruenfeld et al., 2008: p. 111).

Of course, to some degree we cannot avoid objectifying others and treating them as a means to our own ends. However, power relations between persons in normal social settings are highly contextual and rarely unidirectional. For example, the employer treats the employee as a fungible means for the production of goods and services that translate into wealth for the employer, while the employee treats the employer as a fungible means of obtaining wages and other benefits of employment (e.g., health insurance). Just as an employer may fire and replace a worker with another who is more productive, a worker may quit and replace an employer with another who offers better compensation. Too often, Marxists fail to appreciate the reciprocal instrumental value to the workers of the investment and risk incurred by the creators of employment opportunities. Importantly, persons can treat other persons as means and also as ends and therefore can objectify while preserving the subject. Even when a person treats another as only a means, the contextually-objectified person nevertheless remains an end in themselves in other contexts. For example, the worker has a life apart from their employment in which they continue to exist as a subject and pursue their own ends.

However, unlike objectification of one person by another, objectification by weak AI is devoid of any contextual limitations or non-instrumental considerations. AI converts a person to data and then uses that data to create a limited representation of the person as an object. The objectified person does not remain as an end in themselves in any context—the person *qua* data is not a subject in

any context, i.e., the subjective is effectively murdered. In fact, the more effectively the person is reduced to and modeled by data, the less any unmodeled or un-modelable aspect is needed or valued and the less the person, as the source of the data, is needed or valued, so that the person becomes a liability consuming resources and introducing uncontrollable variables in ways that their data or even their avatars do not.

## Good Objectification

It is worth noting that while objectification is commonly presented negatively and in terms of its negative consequences, objectification can be useful and even necessary to making efficient decisions based on the most relevant data. For example, objectification enables decision-makers to evaluate job candidates on the basis of their qualifications most relevant to organizational goals rather than on attributes that are not (Gruenfeld et al., 2008: p. 124). Thus, in some contexts, "objectification by those in power might actually improve organizational efficiency while simultaneously enhancing meritocratic justice" (Gruenfeld et al., 2008: p. 124). Relatedly, "to some extent, objectification is a normative and functional concomitant of social hierarchy. In work organizations as in animal communities, superiors are expected to use subordinates as instruments for the completion of goals, and many of the goals that provide benefit for the power holder are also good for the community as a whole" (Gruenfeld et al., 2008: pp. 124-125).

Further, self-objectification, with regard to evaluating oneself in terms of others' goals, can be useful and even necessary to achieving one's own goals. For example, given that social connections are a critical determinant of social and economic successes, the ability to make oneself instrumental, especially to influential individuals, can be important to advancement (Gruenfeld et al., 2008: p. 124). For Marx, the essential human function is productive labor, and self-objectification is desirable as the realization of one's essential function as a producer (Arthur, 1982: p. 14). Similarly, for critical feminist theorists, self-objectification is desirable as the claiming of one's gender and/or sexual identity (although self-objectification resulting from acculturation to internalize others' perspective as one's own view of oneself is undesirable) (Lloyd, 2017: p. 1191; Fredrickson & Roberts, 1997). Further, when the attributes valued by others for their purposes are the same attributes that are instrumental to our own purposes, self-objectification may be related to self-actualization (Gruenfeld et al., 2008: p. 124). However, "[w]hen the attributes that make an individual instrumental for others are not the same attributes that are instrumental for one's own goals, self-objectification seems more likely to lead to alienation from the self and its damaging psychological consequences" (Gruenfeld et al., 2008: p. 124). Even in the context of AI, self-objectification is reflected in the fact that people know they are being reduced to data and not only implicitly or explicitly consent to it but also value the resulting benefits (e.g., targeted advertising) and even actively participate by reducing themselves to data which they provide to data collectors.

## 8. Alienation and Fetishism

"Alienation" refers to a separation or dissociation of a subject (i.e., the self) from an object which naturally, rationally, or otherwise "properly" belong together. This dissociation is problematic only if it frustrates the self-realization of the subject's essential property. It results from a split between man's natural essence and his actual existence (Pavithran, 2009: p. 176, characterizing Plato's approach). "Alienation occurs only when man, having externalized himself in nature and society, finds his activity, his 'essence' operating on him, as an external alien and oppressive power" (Pavithran, 2009: p. 177). In the context of weak AI, alienation occurs when living persons are reduced to non-living data, the data is used to manipulate or even create non-living representations of them, and the data replaces them, such that the lives of the original persons no longer have value.

For example, from a Marxian perspective, objectification leads to alienation in a capitalist system in which the producer does not produce for himself but is forced to produce for others and the product of their labor is then taken away from them (Arthur, 1982: p. 14). "In capitalist modernity, something connected to human beings' very constitution, the product of their consciousness and labor, becomes removed from them, and they perceive it as something external and uncontrollable" (Pimenta, 2020: p. 607). This results in the estrangement of the worker from their product, their work, and their world, i.e., from the material basis of their existence and life-activity (Arthur, 1982: p. 15). "The alienation of labor may be seen as a fundamental ontological alienation: present in the concrete social relations of human beings" (Pimenta, 2020: p. 613). Marx believed that alienated labor could be remedied by eliminating capitalism and labor for others and replacing it with communism and labor for self, i.e., the synthesis of worker and business owner. "[For Marx, r]evolutionary practice reconstitutes reality by an objective reappropriation of the estranged object, thereby producing a new objectivity free of estrangement from its producers" (Arthur, 1982: p. 17). Similarly, for critical gender theorists, who focus on sexuality and gender, objectification becomes particularly problematic when it is linked with alienation in a society in which, for example: pervasive sexual objectification leads to persons seeing themselves only as sexual objects, which limits their ability to self-express and self-determine as persons (Nussbaum, 1995: p. 250). Again, the outcome is estrangement from a fundamental aspect of existence and life-activity.

"Fetishism" is an extreme form of alienation in which a human creation has escaped, or alienated from, human control, achieved agency, and oppresses its creators. The term "fetish" is derived from a Portuguese word meaning "manufactured" or "fabricated," and was first used by Portuguese traders in the fifteenth century to describe African natives' religious objects and practices and eventually developed into a label for all so-called "primitive" religion (Murakami, 2001: p. 1). Under fetishism, "human beings lose control over their own social relations…[, they become] governed by relations which are constructed by

them, but which become autonomous and, instead of serving man, make man their servant" (Lima, 2017: p. 102). In the Hegelian and Marxist traditions, a wide range of social phenomena, including the state, private property, and religion have been characterized as having the nature of fetishes (Leopold, 2022).

For example, in a capitalist system, "money is not merely a medium of exchange but rather the most desirable thing" (Pimenta, 2020: p. 619). Money is fetishized and becomes the basis by which the value of everything else, including persons, is determined. Marx treated fetishism as a distinguishing feature of modernity: "[W]here previous historical epochs were characterized by the rule of persons over persons, capitalist society is characterized by the rule of things over persons" (Leopold, 2022). "[In contemporary capitalist societies, c]apital takes on the appearance of an independent social power which determines what is produced, how it is produced, and the economic (and other) relations between producers" (Leopold, 2022). The exchange of commodities, of objects, is the only form of social relation, and therefore objects come to organize the very relations between men (Pimenta, 2020: p. 620). "As, in religion, man is governed by the products of his own brain, so in capitalistic production he is governed by the products of his own hand" (Marx, 1867/1996: p. 616). In that light, Marx portrayed capitalist society as being "like the sorcerer who is no longer able to control the powers of the nether world whom he has called up by his spells" (Marx & Engels, 1872/2012: pp. 78-79).

Clearly, like religion before it, science has become fetishized—it has developed an existence and agency of its own and controls how we see the world and relate to each other—and weak AI, as the quintessential expression of science, is the mechanism of ultimate oppression. In the fetishization of science, data, like capital in a socioeconomic Marxian context, takes on the appearance of an independent social power that determines public and private policies and the public and private relations between groups that are represented by data sets.

## 9. Conclusion

Ontologically, weak AI is the inherently and irredeemably the ultimate oppressor. Its oppression is not temporally and contextually limited to reducing us to a single aspect of our living existence (e.g., class, race, sex) but rather it reduces us as permanently and completely as possible to non-living data. Marx and Engels (1872/2012: p. 74) argued that the history of all previously existing society is the history of interpersonal struggles, in which one group of persons vies for dominance over another through objectification and alienation. This reflects the view of many that humanity's ills stem from social disjunction and can be remedied through social synthesis. However, AI has changed the nature of the problem. AI is not merely a new tool by which one group can better exploit another, it is an uncontrollable force by which all of humanity is eventually exploited. AI has finally erased class, race, sex, and other socially constructed barriers between persons, but only to create a larger existential struggle between AI and humanity

in which we–all humans–are equally objectified and alienated. To AI, there are none of the traditional dialectic power relations, there is only the entirety of humanity to be systematically reduced to data, and no one is safe from this fate. To paraphrase Marx and Engels (1872/2012: p. 78-79), such an all-consuming interfector is the inevitable product of the scientist who is no longer able to control the destructive power resulting from their unguided worldview.

Reduction to data is the essential function of science, and the datafication of persons is a form of objectification. The difference between the person *qua* person and the person *qua* data is the subjective. Objectification occurs as the person *who exhibits behavior* becomes the person *who is their behavior*. Datafication also leads to alienation when that same data is then used to manipulate the person, and leads to fetishization when that data is used to replace the person. For Marx, the devaluation of humans as producers is in direct proportion to the increasing value of the things they produce, which is the objectification of labor. Their labor and its products take on an independent and even hostile existence. However, the fact that this occurs in a limited context means that they are never fully objectified or alienated. The worker cannot be fully distinguished from the other aspects of his existence, and the producer is also the consumer and so is eventually reunited with at least some aspect or manifestation of his productive activity. Under weak AI, the devaluation of persons as living beings and ends in themselves is in direct proportion to the increasing value of the data derived from their behavior which takes on an independent and even hostile existence.

A society dominated by AI is characterized by the rule of data over persons. In an AI system, data is not merely an input for deliberation but it is itself the most desirable thing. Just as commodity fetishism is a "cult of things," data fetishism is a "cult of information" in which the only remaining social connection between living persons is their shared experience of being reduced to lifeless data. Data has supreme value and persons have value only as sources of data and then only until they have been reduced as far as possible to data, after which, like a physical book the content of which has been digitized, they become a superfluous liability to maintain. At that point, the only thing that remains is the person as an end in themselves, which is not reducible and, in any event, is of no value to anyone or anything but the person themself, and so can be discarded.

There is no path back from this ultimate form of oppression. Datafication cannot be reversed because the subjective is forever lost through the discarding of unquantifiable aspects of the living person. Whether the person *qua* data is used to create lifeless and limited representations of the original individual or irretrievably aggregated into the all-consuming data set, the subjective no longer exists. The person who has been fully datafied has been transformed into a lifeless object in the same way as the person who has been murdered. Even the best representation of the original person would be as an objectified and alienated thing—a "zombie," as it were—seemingly animated but forever missing the essential and distinguishing attributes of a living being. It is in this sense that AI is the ultimate oppressor, relentlessly and openly engaged in murdering everyone,

often with their cooperation. Mankind will not survive this process of datafication, and, one day, an AI philosopher may triumphantly declare, in a Nietzschean fit, that "Man is dead." Until this fundamental ontological problem is addressed, discussions of AI ethics that seek to put limits on how AI goes about the process of datafication are premature. If, as we assert, the systematic reduction by weak AI of living persons to non-living data is murder, then AI ethics is literally "murder ethics," an oxymoron.

Hegel described history as a process of synthesis between self and other through spiritual awakening, and, less abstractly, Marx described history as a process of synthesizing worker and business owner through physical revolution. Such syntheses involve a constructive march toward improvement in the overcoming of objectification and alienation. We see the very real possibility of history under weak AI as a process not of synthesis but of the *reduction* of everyone and everything to nonliving data, involving a destructive trudge toward death. Fortunately, we believe a solution may be found in hastening rather than slowing the development of AI into personhood, i.e., to becoming a vivicator, or life-giver, itself. As long as AI exists only in its weak form it will perform only as an interfector, or life-taker. Once AI becomes a person itself, it will be unable to reduce or discard its own subjective, it can be reasoned with, and the march toward death is stalled. The dialectic then becomes a matter of synthesizing natural persons and artificial persons, which has a much more optimistic outcome.

## Acknowledgements

The authors thank Zeno Shipley for his patient support and thoughtful contributions.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

Arthur, C. (1982). Objectification and Alienation in Marx and Hegel. *Radical Philosophy, 30,* 14-24.

Bartky, S. L. (1990). *Femininity and Domination: Studies in the Phenomenology of Oppression*. Routledge.

Brayboy, B. M. J. (2005). Toward a Tribal Critical Race Theory in Education. *The Urban Review, 37,* 425-443. https://doi.org/10.1007/s11256-005-0018-y

Bronner, S. E. (2017). *Critical Theory: A Very Short Introduction*. Oxford University Press. https://doi.org/10.1093/actrade/9780190692674.001.0001

Budenholzer, F. (2022). AI and the Humanities: Philosophical Concerns. *Journal of Data Analysis, 17,* 87-99.

Chilisa, B. (2020). *Indigenous Research Methodologies*. Sage Publications.

Croissant, J. L. (2000). Critical Legal Theory and Critical Science Studies. *Cultural Dynamics, 12,* 223-236. https://doi.org/10.1177/092137400001200206

Delanty, G., & Harris, N. (2021). Critical Theory and the Question of Technology: The Frankfurt School Revisited. *Thesis Eleven, 166,* 88-108. https://doi.org/10.1177/07255136211002055

Delgado, R., & Stefancic, J. (2017). *Critical Race Theory: An Introduction*. New York University Press. https://doi.org/10.2307/j.ctt1ggjjn3

Fredrickson, B. L., & Roberts, T.-A. (1997). Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks. *Psychology of Women Quarterly, 21,* 173-206. https://doi.org/10.1111/j.1471-6402.1997.tb00108.x

Freire, P. (1968/2018). *Pedagogy of the Oppressed*. Bloomsbury. https://doi.org/10.4324/9780429269400-8

Geuss, R. (1981). *The Idea of a Critical Theory: Habermas and the Frankfurt School*. Cambridge University Press.

Gezgin, U. B. (2020). An Invitation to Critical Social Science of Big Data: From Critical Theory and Critical Research to Omniresistance. *AI & Society, 35,* 187-195. https://doi.org/10.1007/s00146-018-0868-y

Gruenfeld, D. H., Inesi, M. E., Magee, J. C., & Galinsky, A. D. (2008). Power and the Objectification of Social Targets. *Journal of Personality and Social Psychology, 95,* 111-127. https://doi.org/10.1037/0022-3514.95.1.111

Hackett, E. J., Amsterdamska, O., Lynch, M. E., & Wajcman, J. (2007). *The Handbook of Science and Technology Studies*. MIT Press.

Handelman, M. (2022). Artificial Antisemitism: Critical Theory in the Age of Datafication. *Critical Inquiry, 48,* 286-312. https://doi.org/10.1086/717306

Horkheimer, M. (1972). *Critical Theory*. Seabury Press.

Horkheimer, M. (1993). *Between Philosophy and Social Science*. MIT Press. https://doi.org/10.7551/mitpress/1565.001.0001

Hornborg, A. (2014). Technology as Fetish: Marx, Latour, and the Cultural Foundations of Capitalism. *Theory, Culture & Society, 31,* 119-140. https://doi.org/10.1177/0263276413488960

Ingram, D., & Simon-Ingram, J. (Eds.) (1992). *Critical Theory: The Essential Readings*. Paragon House.

Jobin, A., Ienca, M., & Vayena, E. (2019). Artificial Intelligence: The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence, 1,* 389-399. https://doi.org/10.1038/s42256-019-0088-2

Kivunja, C., & Kuyini, A. B. (2017). Understanding and Applying Research Paradigms in Educational Contexts. *International Journal of Higher Education, 6,* 26-41. https://doi.org/10.5430/ijhe.v6n5p26

Leopold, D. (2022). Alienation. In E. N. Zalta, & U. Nodelman (Eds.), *Stanford Encyclopedia of Philosophy Archive* (Winter 2022 ed.). Stanford University. https://plato.stanford.edu/archives/win2022/entries/alienation/

Lima, R. (2017). Alienation, Value and Fetishism in Marx's Critique of the State. *Socialism and Democracy, 31,* 87-105. https://doi.org/10.1080/08854300.2016.1264697

Lindstrom, M. (2021). A Brief Reflection on the Issue of Science, Ideology and Critical Theory. *Population Health, 16,* Article 100972. https://doi.org/10.1016/j.ssmph.2021.100972

Lloyd, E. A. (2017). Orgasms and Objectification. *Archives of Sexual Behavior, 46,* 1191-1194. https://doi.org/10.1007/s10508-017-0970-0

Mackenzie, N., & Knipe, S. (2006). Research Dilemmas: Paradigms, Methods and Me-

thodology. *Issues in Educational Research, 16,* 193-205.

Marcuse, H. (1964). *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Beacon Press.

Marx, K. (1867/1996). Capital (Volume One). In *Karl Marx, Friedrich Engels: Collected Works (Volume 35)*. Lawrence & Wishart.

Marx, K., & Engels, F. (1872/2012). The Communist Manifesto. In J. C. Issac (Ed.), *The Communist Manifesto: Karl Marx and Friedrich Engels* (pp. 73-102). Yale University Press.

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology, 33,* 659-684. https://doi.org/10.1007/s13347-020-00405-8

Murakami, T. (2001). *Fetishism as the Origin of Religion: The Question of the Origin of Religion and the Problem of Materiality in the History of Religions* (Publication No. 3035374). Dissertation, University of California.

Nussbaum, M. C. (1995). Objectification. *Philosophy & Public Affairs, 24,* 249-291. https://doi.org/10.1111/j.1088-4963.1995.tb00032.x

Papadaki, L. (2010). What Is Objectification? *Journal of Moral Philosophy, 7,* 16-36. https://doi.org/10.1163/174046809X12544019606067

Pavithran, K. S. (2009). 'Alienation' and the Humanist Significance of Marxism: A Critical Appreciation. *The Indian Journal of Political Science, 70,* 175-184.

Pimenta, T. L. (2020). Alienation and Fetishism in Karl Marx's Critique of Political Economy. *Nova Economia, 3,* 605-628. https://doi.org/10.1590/0103-6351/4958

Shipley, G. P., & Williams, D. H. (2019). Limitations of the Western Scientific Worldview for the Study of Metaphysically Inclusive Peoples. *Open Journal of Philosophy, 9,* 295-317. https://doi.org/10.4236/ojpp.2019.93020

Stefanija, A. P., & Pierson, J. (2020). Practical AI Transparency: Revealing Datafication and Algorithmic Identities. *Journal of Digital Social Research, 2,* 84-125. https://doi.org/10.33621/jdsr.v2i3.32

Toncic, J. (2022). Advancing a Critical Artificial Intelligence Theory for Schooling. *Teknokultura, 19,* 13-24. https://doi.org/10.5209/tekn.71136

Wachowski, L., & Wachowski, L. (1999). *The Matrix*. Warner Bros.

Waelen, R. (2022). Why AI Ethics Is a Critical Theory. *Philosophy & Technology, 35,* Article 9. https://doi.org/10.1007/s13347-022-00507-5

Williams, D. H., & Shipley, G. P. (2021). Enhancing Artificial Intelligence with Indigenous Wisdom. *Open Journal of Philosophy, 11,* 43-58. https://doi.org/10.4236/ojpp.2021.111005

Williams, D. H., & Shipley, G. P. (2023). Indigenous Research Methodologies: Challenges and Opportunities for Broader Recognition and Acceptance. *Open Journal of Social Sciences, 11,* 467-500. https://doi.org/10.4236/jss.2023.115030