Journal of Posthumanism
Fall 2024
Volume: 4, No: 3, pp. 231–245
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
journals.tplondon.com/jp

TRANSNATIONAL PRESS® LONDON

Received: 5 March 2024 Accepted: 19 September 2024
DOI: https://doi.org/10.33182/joph.v4i3.3287

Towards a Posthumanist Critique of Large Language Models

Claudio Celis Bueno[1] and Jernej Markelj[2]

Abstract

This article develops a critique of large language models (LLMs) from a posthumanist perspective. The first part focuses on Emily Bender's critique of LLMs in order to highlight how its conceptual and political axioms have informed recent critiques of ChatGPT. We make a case that this anthropocentric perspective remains insufficient for adequately grasping its conceptual and political consequences. In the second part of the article, we address these shortcomings by proposing a posthumanist critique of LLMs. To formulate this critique, we begin by drawing on Eric Hörl's contention that the age of digitalization (what he calls "cybernetization") demands a radical redefinition of the concept of "critique" (Hörl et al., 2021, 7). Relying on Hörl's intervention, we then gradually develop a posthumanist framework by grounding it in four interlinked concepts: general ecology, machinic agency, machinic surplus value, and cosmotechnics. After advancing the said theoretical framework, our conclusion mobilises it to outline a posthumanist critique of LLMs.

Keywords: *Critical Posthumanism; Machinic Agency; Large Language Models; Anthropocentrism; General Ecology*

Introduction

Since its public release in November 2022, *ChatGPT* has triggered a significant amount of both enthusiastic and critical responses. A substantial part of the critical responses focuses on the disruptive effects that this technology is having on existing practices (e.g., education, journalism, creative labour, computer programming, etc.). Other responses focus on the economic and environmental toll that this technology might entail. Finally, more conceptual critiques highlight the philosophical, societal, and ethical implications of large language models being used in a systematic and widespread manner.[3] Yet, it is important to note that the systematic critique of large language models (LLMs) precedes the release of *ChatGPT* and the current hype around "generative AI".

* This article was published through an open-access model that charged no article processing fees.

[1] Claudio Celis Bueno, University of Amsterdam. E-mail: c.o.celisbueno@uva.nl

[2] Jernej Markelj, University of Amsterdam. E-mail: j.markelj@uva.nl

[3] In this third group we find a range of arguments. Some authors criticize *ChatGPT* for its inability to infer causal explanations (Chomsky et al., 2023; Chiang, 2023). These authors stress that *ChatGPT* is limited to identifying statistical correlations between words, and incapable of abstracting from them any kind of causal laws or principles (like, e.g., the basic principles of algebra). Other critiques have been levelled against *ChatGPT* for its inability for critical and contextual thinking, but also for its lack of moral thinking (Chomsky et al., 2023). The latter is a matter of the capacity to constrain "the otherwise limitless creativity of our minds with a set of ethical principles that determines what ought and ought not to be" (Chomsky et al., 2023). There are also critiques that highlight that LLMs "black-box" their training data (Burgess, 2023); that they reproduce socio-political biases (Bender et al., 2021; Motoki et al., 2024); disseminate misinformation (Bell, 2023; Bridle, 2023; De Angelis et al., 2023; Hsu and Thompson, 2023); or on account of their "hallucinations", the non-sensical or counterfactual statements that *ChatGPT* and other LLMs occasionally produce (Chiang, 2023; Emsley, 2023; Metz, 2023). Finally, large language models have been also criticized for their huge environmental costs (Bender et al., 2021; Bridle, 2023), the precarious ghost labour behind their training process (Perrigo, 2023), and the appropriation of collectively produced (and likely copyrighted) texts for its training purposes (Bode and Goodlad, 2023).

Several of the concerns outlined above have been most notably discussed by computational linguist Emily Bender (2022; 2024; Bender and Koller, 2020; Bender et al., 2021). For Bender, the most prominent risk posed by LLMs stems from our tendency to anthropomorphise their ability to process natural language. While LLMs are mere probabilistic machines without access to "meaning", she observes that we tend to imagine a mind behind their output and, thus, treat this output as "meaningful". By anthropomorphising LLMs, Bender (2024) suggests, we risk "dehumanising" our particular relation to language and, with it, undermining human politics and ethics. To avoid this, the line that separates human from machine language must be alertly policed. As we show below, Bender's anthropocentric position informs many of the recent critical responses to LLMs, in particular since the release of *ChatGPT* (Chomsky et al., 2023; Bridle, 2023; Chiang, 2023; Gupta et al., 2024; Metz, 2023; Weil, 2023).

Building on critical engagement with Bender's work, this article argues that most of these recent critiques presuppose conceptual and political axioms that (sometimes inadvertently) reinforce an anthropocentric understanding of LLMs. We make a case that this anthropocentric perspective, which frames human-technology interactions as instrumental relations between autonomous subjects and essentially passive and unintelligent technologies, remains insufficient for adequately grasping its conceptual and political consequences. We address these shortcomings by proposing a posthumanist critique of LLMs. In line with the critical posthumanist approach proposed by Sylvia IV (2021), our critique proceeds by shifting focus from human subjects understood as stable, unified and self-determined, to wider socio-technical systems that constantly (re)produce these subjects. To formulate this critique we begin by drawing on Eric Hörl's contention that the age of digitalization (what he calls "cybernetization") demands a radical redefinition of the concept of "critique" (Hörl et al., 2021, 7). Relying on Hörl's intervention, we then gradually develop a posthumanist framework by grounding it in four interlinked concepts: general ecology, machinic agency, machinic surplus value, and cosmotechnics.[4] After advancing the said theoretical framework, our conclusion mobilises it to outline a posthumanist critique of LLMs.

Anthropocentric critiques of anthropomorphised technologies

Warnings about the detrimental effects of anthropomorphising the output of statistical machines issued by Bender echo in recent critiques of LLMs. Still, concerns that draw attention to the distinction between human understanding and the ways in which machines process language have been raised before. An early version of this critique can be traced back to John Searle's (1980) famous "Chinese Room" experiment. Searle imagined an English speaker being placed in a room and provided with a set of instructions (in English) detailing how to process Chinese symbols written on cards and produce appropriate responses based on those symbols. When prompted with Chinese utterances, this person can produce coherent answers even though they do not understand Chinese. With this thought experiment, Searle sought to challenge the claim that computers exhibit true human-like intelligence and understanding. In his view, computers can produce seemingly meaningful statements and create an impression of understanding, but they do so only by blindly manipulating symbols.[5]

[4] We define "posthumanist framework" in a twofold sense: a) by following Hörl's (2021) call for a radical redefinition of critique in the current context of digitalisation; and b) as "critical posthumanism", a "theoretical approach" developed by authors such as Rosi Braidotti and Katherine Hayles aimed at the "deconstruction of humanism" (Herbrechter, 2018, 94).

[5] Searle (1980) distinguishes between a "syntactic" and a "semantic" relation to language. According to him, when computers are trained to perform the former, we can speak of "weak AI". Only in an eventual scenario in which computers were able to establish a semantic relation to language could we speak of a "strong AI", or "human-like intelligence".

Unlike Searle, whose critique of computer simulation is grounded in analytic philosophy, Bender advances her analogous critique of LLMs from the perspective of computational linguistics. Together with other scholars (Bender and Koller, 2020; Bender et al., 2021), she suggests that LLMs are nothing more than "stochastic parrots" as they operate by means of probabilistic mimicking of previously existing texts. Bender's argument is based on establishing a distinction between linguistic form and meaning (2020, 5186). While Bender takes linguistic form to be any symbolic manifestation of language, be it "marks on a page, pixels or bytes in a digital representation of text, or movements of the articulators", she defines meaning as a relation between linguistic form and the communicative intent of a speaker (2020, 5186-7). For her, understanding the meaning of language corresponds to grasping the communicative intent behind the linguistic form, which is achieved by connecting the latter to "objects outside of language", that is, to "the speakers' actual (physical, social, and mental) world" (Bender and Koller, 2020, 5188). Since language models do not have access to these external referents, but are instead trained on, and operate with, exclusively linguistic form, they are incapable of relating linguistic signs to their referents, which is in Bender's view the very connection that constitutes meaning.

Bender thus sees LLMs as probabilistic machines that mindlessly produce texts by rearranging their training data. Still, the main issue for her is that we tend to interpret these texts as though there was a communicative intent behind them. Due to the perceived fluency and coherence of these texts, we habitually consider their conventional meaning and their context to construct the mental state of an interlocutor (their thoughts, intentions, credibility).[6] Put differently, by interpreting LLMs' responses as if they were human or human-like Bender suggests that we anthropomorphize them. This habitual tendency, coupled with the chatbot's design that deliberately encourages it by mimicking humans (Bender in Weil, 2023), bring about different forms of risks and dangers (Bender et al., 2021). At the more extreme end, Bender warns that anthropomorphising LLMs can lead to the dehumanization of humans (Bender, 2024). She defines this dehumanization as "(a) the cognitive state of failing to perceive another human as fully human, (b) an act that expresses that cognitive state […], or (c) the experience of being subjected to an act that expresses a lack of perception of one's humanity and/or denies human experience or human rights, or combinations thereof" (Bender, 2024, 115). While Bender is critical of our tendency to anthropomorphise LLMs, her critique ultimately aims at reinforcing the boundary between humans and machines. In her view, this boundary needs to be policed to ensure that politics do not become pure technical calculation and that humans do not end up being denied their "human experience or human rights".[7]

Bender's claim that LLMs are incapable of grasping the meaning of language is echoed by a series of recent critiques of LLMs, particularly of OpenAI's *ChatGPT*. James Bridle (2023), for instance, sees *ChatGPT* as "inherently stupid" and actively dangerous. Like Bender, Bridle grounds his claims about LLMs' stupidity in its inability to connect words to their reality. He proposes that *ChatGPT* "has read most of the internet, and it knows what human language is supposed to sound like, but it has no relation to reality whatsoever" (2023). Since it is capable of creating the appearance of meaning, Bridle fears that it will be taken as a reliable source of information. Comparable claims about *ChatGPT*'s stupidity are made by Ian Bogost (2023). In his view,

[6] Similarly, Gupta et al. (2024) suggest that the metaphors we use in relation to ChatGPT can also contribute to its anthropomorphizing.
[7] Bender (2024) identifies six ways in which this dehumanization can take place: a) using computational metaphors that equate the brain with a computer and a computer with the brain; b) digital physiognomy; c) ignoring the human (hidden)labour behind AI; d) the belief that datasets are "representative"; e) the "irrelationality" of AI models; f) and the reinforcement of the "white racial frame".

> *ChatGPT* lacks the ability to truly understand the complexity of human language and conversation. It is simply trained to generate words based on a given input, but it does not have the ability to truly comprehend the meaning behind those words.

In line with Bender and Bridle, Bogost emphasizes the probabilistic nature of texts generated by *ChatGPT*. Since the chatbot has no access to the embodied referents of language, he observes that the generated utterances lack in depth and insight. Corresponding limitations of *ChatGPT*, when compared to a human mind, are glossed by Chomsky et al. (2023) and Ted Chiang (2023). While the former argue that LLMs "differ profoundly from how humans use reason and language", the latter differentiates between the lossless compression of human understanding and the lossy text-compression performed by *ChatGPT*.

According to some (Coeckelbergh and Gunkel, 2023; Tufecki, 2022), the argument made by Bender and others can be traced all the way back to Plato's *Phaedrus*. There Plato condemns the sophistic practice of teaching with the aid of writing. For him, learning from written texts (as opposed to *via* dialogues with a teacher) produces false knowledge, "the conceit of wisdom" (Plato, 1997, 157). This false wisdom corresponds to abstract, non-assimilated knowledge. "By telling them of many things without teaching them", suggests Plato (1997, 157), writing "will make [students] seem to know much, while for the most part they know nothing". In his view, the knowledge contained in, and disseminated by, the written word is deceptive as it is not substantiated by lived experience. Plato adds that written texts "seem to talk to you as though they were intelligent, but if you ask them anything about what they say, […] they go on telling you just the same thing for ever" (1997, 158). To the mute symbols on a page and the defective ideas they engender, Plato (1997, 159) opposes knowledge that "is written in the soul of the learner" and is transmitted by means of living speech. "It is speech", suggest Coeckelbergh and Gunkel (2023),

> and its connection to the living voice of the speaker—the embodied human being who lives in the world and know what it is they speak about—that authorizes and guarantees the truth of what is said. Speech has a direct and intimate connection to the real.

Unlike the written texts condemned by Plato, LLMs do have the ability to "respond". At the same time, however, it can be said that the logic of Plato's critique of writing is still aligned with the critique of LLMs articulated by Bender and others. Firstly, both critiques are grounded in the distinction between mere symbols on the one hand, and the immediacy of the living voice and embodied experience on the other. Moreover, they both define the former as inferior to the latter. This hierarchical distinction between the embodied mind (and speech as its pure and immediate expression) as the privileged site for accessing meaning and writing as its derivative technological representation is what Jacques Derrida (1976) refers to as "logocentrism". For him, logocentrism is in fact a tendency that underlies the entire tradition of Western science and philosophy.

Derrida (1976) famously deconstructs the hierarchical binary that grounds logocentrism. In short, he argues that writing is not merely an artificial and corrupting supplement that is secondary to speech. Instead, Derrida suggests that writing is what shapes the subjectivity of a speaker in the first place. Thus, writing not only precedes living speech, but actually makes it possible. Derrida's critique of logocentrism is further developed by Bernard Stiegler (1998), who extends this critique from writing to technology in general. According to Stiegler, human subjects are characterized by "originary technicity" (1998): we are not autonomous agents fully in control of our external technological prostheses, but instead animals that have invented ourselves as humans only through the use of technologies. If writing does not merely exteriorize our pre-existing thoughts but is a

condition of possibility for their constitution, the same constitutive relation applies to every other technology that we interact with as they too, for better or worse, shape our sensory, cognitive, and affective capacities.

Coeckelbergh and Gunkel (2023) suggest that the critique of logocentrism has significant consequences for understanding LLMs. In particular, they claim that this critique "undermines the very notions of authority, authorship, and responsibility" (Coeckelbergh and Gunkel, 2023). If human subjects are constituted and shaped by their interaction with technology, the idea of an autonomous author, who is the sole source of their words, and can be held responsible for them, is no longer tenable. Instead, authorship becomes distributed between humans and non-humans, who can be said to co-produce textual outputs.

This critique of the anthropocentrism of Bender's approach is also advanced by N. Katherine Hayles (2022). She draws attention to its anthropocentric bias and suggests that the claim that LLMs are nothing but a stochastic parrot requires "an implicit assumption that human cognition is the only cognition that really counts" (2022, 652). Unlike Bender, who sees humans alone as capable of using language and thinking, Hayles maintains that "parrots—like all life forms—also have cognitive capabilities, as do large language models such as GPT-3" (2022, 652-3). Instead of dismissing LLMs as mindless and incapable of adequately possessing language, Hayles (2022, 647) seeks to understand their non-human intelligence by reconstructing what constitutes its 'umwelt' (its algorithmic architecture, training input, functioning, etc.).[8]

Another critique of Bender's anthropocentricism is articulated by Tobias Rees (2022), who historicizes the concept of meaning. While Bender presents her definition of meaning as universal, timeless, and hence inevitable, Rees situates it in a specific epistemic paradigm that first emerged in the early twentieth century. In the wake of enlightenment and industrial revolution, and the rise of science and individuality, this paradigm sees language as the human ability to assign and navigate meaning in a meaningless world. According to Rees, the rise of LLMs disrupts this current epistemic paradigm and allows us to formulate more productive ways of thinking about language and intelligence, and the relation between humans and machines.[9] In his view, this is possible as the emergence of LLMs provides us with a practical embodiment of a structuralist theory of language (as advanced by de Saussure, and further developed by others), which frames it as a combinatorial system that functions independently of (the communicative intent of) human subjects. "The power of this new concept of language that emerges from GPT-3", writes Rees (2022, 180), "is that it disrupts human exceptionalism". It does so by undermining the idea of language (but also of thought and intelligence) as something exclusive to humans and makes way for extending them to the domain of animals, microbes, and machines. From the perspective of this developing epistemic paradigm, human intelligence would differ from the intelligence of the latter only in degree, and not in kind. Following these critiques of the anthropocentrism of Bender's framework, the remainder of this article is aimed at setting the ground for a posthumanist critique of LLMs.

[8] By focusing on the "umwelt" of LLMs, their world-horizon, Hayles (2022; 2023) seeks to investigate the interactions between the components that shape and enable this "non-human intelligence". As it zooms in on the enabling (and disabling) connections that constitute LLMs' ecology, her approach is similar to the posthumanist approach that we propose in this article. In fact, one important link between Hayles' approach and some of the Deleuzian concepts deployed below can be found in the notion of umwelt developed by Jakob von Uexküll (see Deleuze, 1988).

[9] Similarly, Weatherby and Justie (2022) argue that we require a new theory of signs in order to fully grasp LLMs. They call this an "indexical" understanding of AI and distinguish it from the more traditional understanding of AI (which they call "symbolic"). From the perspective of Weatherby and Justie, it could be argued that Bender's critique of LLMs reproduces a symbolic notion of language.

Towards a posthumanist critique of digital technologies

To move towards a posthumanist critique of LLMs we follow Eric Hörl's suggestion that the current context of digitalisation demands a radical shift in "the meaning of critique" (Hörl et al., 2021, 7). This shift entails a twofold redefinition of the notion of critique. The first part of this redefinition concerns critique in a Kantian sense: the analysis of the conditions of possibility of a given phenomenon ("transcendental critique"). For Kant, the object of critique is the "transcendental subject", locus of all legitimate knowledge and experience (Hörl et al., 2021, 9). The general process of digitalisation, Hörl et al. (2021, 9) argue, has decentred the "transcendental subject" as the "central actor" and replaced it with the notion of the "environment". As we show in the following subsection, environment for Hörl does not refer to the "natural environment" (a concept that still requires the anthropocentric opposition between nature and culture), but rather to the result of a process of "environmentalization" which blurs the culture-nature-technics divide that is characteristic of modern critique.

The second part of the redefinition of critique refers to critique as a practice (an "ethos" or "attitude") characteristic of modernity. The central question posed by this second form of critique has been the following: what are the "structures" through which a given phenomenon appears as if it were "universal and necessary"? (Hörl et al., 2021, 7). Conversely, Hörl suggests that the key question any "critical project" should ask today is this: to what extent does the passage from the "transcendental subject" to "environmentality" as the condition of possibility of knowledge and experience demand new concepts to analyse the reproduction of power relations in the contemporary world?

According to Hörl et al. (2021, 8), the current context of digitalisation confronts us with the task of an "urgent reproblematization of critique under the conditions of digitality", which requires reformulating the concept of critique in relation to these two aspects. In what follows we present four concepts that, we claim, are instrumental in formulating this reproblematization of critique in the current conditions of digitalisation. In doing so we seek to shift the critique of LLMs from an anthropocentric framework (as adopted by Bender and others) towards a posthumanist and post-anthropocentric one.

A. General Ecology

In recent years several authors have raised the issue of the environmental footprint of machine learning (Bender et al., 2021; Crawford, 2021; Hao, 2019; Kanungo, 2023; Ligozat et al., 2022; Perucica and Andjelkovic, 2022; Valdivia, 2022). These critiques often highlight the high energy consumption of training processes, the powering and cooling of data storage for training datasets, the negative impact of the mining industry required for the production of the digital components, and the digital waste produced by increased demand for faster processes and larger storage. Most of these critiques, however, reproduce anthropocentric notions of environment and ecology that continue to oppose nature to culture. As such, these arguments are insufficient for a posthumanist critique of these technologies.[10] Following Eric Hörl, we argue that the concept of "general ecology" (2013; 2017; 2021) offers a key starting point for a post-anthropocentric critique of LLMs that goes beyond the traditional dichotomy between nature and culture.

[10] Coeckelbergh (2022) explores a potential "posthumanist" critique of AI by focusing on "environmental politics". Still, his analysis does not seem to move beyond a definition of ecology grounded in the nature-culture divide.

The "general ecology", Hörl tells us, is a new "image of thought", a new "historical semantics", which is "critical of all anthropocentrism" (2017). This new image of thought describes the world from the perspective of a "radical relation" (a "machinic collaboration") between "human and nonhuman agents and forces" (Hörl, 2018, 174). As such, the "general ecology" emerges as a response to the changes brought forward by the growing digitalisation (or "cybernetization") of society since the 1950s. In consequence, Hörl (2018, 174) argues, the general ecology is a central concept for understanding "our posthuman situation".[11]

The concept of general ecology involves both a proliferation and a redefinition of the concept of ecology (Hörl, 2017).[12] In this sense, it puts forth a "denaturalisation" of the concept of ecology, an "undoing of the sutures" that link this concept to an anthropocentric definition of nature (2017, 2). Hörl speaks of the current context of digitalisation as a process of "cybernetization". He argues that this new historical semantics began with Cybernetic Theory. As such, the concept of general ecology results from a project of "regulation and control" (2017, 3). After cybernetics, he contends, a serious "revaluation of the sense of ecology" is needed (2013, 128). Hence, instead of speaking of a "human condition" that establishes a sharp separation between humans, nature, and technics, Hörl invites us to speak of a "techno-ecological condition" (2017, 2), in which digital technologies are redefining the boundaries between these three domains.

Hörl suggests that the concept of general ecology responds to a broader turn in theory that can be referred to as the "relational turn" (2017, 6). The relational turn, he tells us, defines relations as "something that precedes the forming of [its] terms (subject, object, individual, groups, indeed all forms of collective human and non-human agents)" (Hörl, 2013, 122). This means that the mediation between an organism and its environment precedes the constitution of the organism and the environment as separate entities. Significantly, the relationality entailed in the general ecology exceeds the domain of human (social) relations and "designates the collaboration of human and nonhuman agents" (2017, 3).

One of Hörl's main contributions towards a critique of "our posthuman situation" has been introducing the concept of "Environmentality" (2017). If the general ecology refers to our new episteme, the new "historical semantics" of our digital age, then Environmentality refers to the changes in the technologies of power that underlie this epistemic transformation. The term Environmentality is deployed as an expansion of Foucault's concept of Governmentality (Foucault, 2009; Hörl, 2018, 175; Hörl, 2017, 49 n.19). In this sense, Environmentality refers to a technology of power that does not focus on the individual subject (as disciplinary power did), but on the regulation and control of the relation between environment and organism. Politics is thus reduced to the regulation of cybernetic environments (involving human and non-human elements). Relations become the object of power, leading to a series of new conceptualisation of politics such as "algorithmic governmentality" (Rouvroy and Berns, 2013), "feed-forward" (Hansen, 2015), "machinic enslavement" (Lazzarato 2014), and "modulation" (Cheney-Lippold, 2011). In the following sub-sections, we flesh out Environmentality as a new technology of power by disentangling the issues of agency, value and technics from their modern, anthropocentric, and universalising frameworks.

[11] Hörl explains the idea of "posthuman situation" by referring to the work of Rosi Braidotti (2013).
[12] Hörl speaks of "a thousand ecologies" as a multiplication of Felix Guattari's (2014) "three ecologies".

B. Machinic Agency

By decentring the transcendental subject as the locus of knowledge and experience and replacing it with the notion of general ecology, the current process of "cybernetization" entails an "explosion of agencies" (Hörl, 2017, 12). This defies the "anthropocentric illusion" characteristic of modernity according to which agency takes places solely on the side of human actors, who use technological means in order to transform nature. The general ecology, instead, acknowledges a proliferation of "environmental agencies" (human and non-human) and thus unveils the "illusionary character of the [human] monopoly on agency" (Hörl, 2017, 12). The proliferation of agencies unleashed by cybernetization corresponds to a series of other approaches to theorising non-human agency, including "agential realism" (Barad, 2007), "material agency" (Bennet, 2009), "deviant agency" (Alaimo, 2010), "distributed agency" (Rammert, 2008), and "environmental agency" (Hansen, 2009).[13] As Marchand (2018, 293) contends, the development of an "expanded notion of agency" capable of accommodating the "multiple non-human 'actants' with whom we share and constitute our common world has become a chief concern among many posthumanist writers".

Within this broad palette of non-human conceptualisations of agency, we argue that the notion of "machinic agency", as conceptualised by Deleuze and Guattari, offers a productive framework for a posthumanist critique of algorithmic technologies. The concept of machinic agency (Markelj and Celis Bueno, 2023) locates agential forces within the productive connections between human and nonhuman entities; it is these affective connections that constitute and exercise their capacities to act. As argued (Markelj and Celis Bueno, 2023), we maintain that the concept of machinic agency is relevant for a critique of contemporary technologies for at least two reasons. First, it allows overcoming the dualism between mechanism and organicism that has informed Western philosophies of technology (Markelj and Celis Bueno, 2023). This is important because this approach breaks away from (a.) the culture-nature divide (that aligns agency with human intentionality); (b.) the nature-technics divide (that aligns agency with living beings); and (c.) the culture-technics divide (that aligns agency with human, non-mechanistic judgment). In this sense, the concept of machinic agency manages not only to blur the distinction between culture, nature, and technics that informs anthropocentrism, but also to bring together the different attempts to conceptualise an expanded (posthumanist) notion of agency mentioned above.

Furthermore, we argue that the notion of machinic agency allows displacing the limits between these three domains while still being able to account for asymmetric relations of power in contemporary capitalism (Markelj and Celis Bueno, 2023). This is an important feature that safeguards the possibility of a critical standpoint. The project of cybernetics, or Environmentality, aims at dissolving the differences between culture, nature, and technics to impose a universal regime of regulation and control based on the relation between organisms and their environments. An approach that blurs these boundaries without accounting for asymmetric power relations risks being complicit with a cybernetic project of total control. Hence, a posthumanist critique of digital technologies must, on the one hand, overcome the anthropocentric definition of agency, but on the other, account for the new power asymmetries that stem from these new forms of control.

C. Machinic Surplus Value

As mentioned above, Hörl defines the second modality of critique as the analysis of the specific structures that, in a given episteme, appear as "universal and necessary" (Hörl et al., 2021, 7). Central

[13] For an overview of a non-human definition of agency see Marchand (2018).

to Marx's (1976) critique of capitalism is his analysis of the "value form". In the terms laid out by Hörl, Marx's analysis belongs to the modern episteme since it requires a sharp distinction between culture, nature, and technics. According to Marx (1976), only humans produce value. This claim is grounded in a sharp differentiation between the intentionality of human labour, the instinctive character of animal behaviour, and the mechanical rules governing technical objects. As the concept of "value form" seeks to unveil the exploitation of human labour, it is at the core of Marx's critique of capitalism. Yet, according to Hörl (2021), an analysis of contemporary capitalism must go beyond this anthropocentric critique of exploitation. He maintains that critique today must focus on how Environmentality (as a new technology of power) begins to operate as an "apparatus of capture" aimed at the "exploitation of relationality" (Hörl, 2021). Put differently, whereas Marx's anthropocentric framework developed a critique of capitalism based on the exploitation of human labour, a posthumanist critique of capitalism must focus on the exploitation of the relationality that characterises the general ecology (Hörl, 2017, 8). While Marx's notion of the "value form" sums up the former, we believe that Deleuze and Guattari's (1983) concept of "machinic surplus value" better illustrates the latter.

Writing in the early 1970s, Deleuze and Guattari (1983) argue that the growing digitalisation of production is causing a shift from a form of exploitation based on human surplus value towards an exploitation of the relationality between different types of machines (human, biological, social, technical, etc.).[14] In light of this, they claim, the Marxist definition of surplus value "must be modified in terms of the machinic surplus value of constant capital, which distinguishes itself from the human surplus value of variable capital" (1983, 239). This does not mean simply that machines become sources of "human-like" surplus value. Rather, it entails a shift in the conceptualisation of surplus value that resonates with the relational turn characteristic of the "general ecology". What becomes the object of capitalist exploitation is a process of "amplification by connection" (Markelj and Celis Bueno, 2023; see also Pasquinelli, 2015). In the context of Environmentality, capitalism becomes an "apparatus of capture" aimed at the capturing of the relationality between an organism and its environment (Hörl, 2021). As such, power ceases to be associated "with ownership of the means of production" and becomes associated with ownership of means of capturing a relation and modifying it (Hörl, 2021, 120).[15]

D. Cosmotechnics

The concept of general ecology leads also towards a non-modern appreciation of the notion of technology (Hörl, 2017, 3). In this sense, Hörl calls for a "non-modern mapping of Environmentality" (2013, 129) and a "perspectivist" approach to technology (2017, 7). The general ecology entails not only the "historical undoing" of Western anthropocentrism but also opens the door for a "plural techno-ecology" (Hörl, 2017, 12).[16] We propose that this shift can be summed up through Yuk Hui's concept of cosmotechnics.

[14] For an overview of the concept of "machinic surplus value" in the current context of digitalization see Celis Bueno (2024).

[15] Hörl (2021) links the idea of an apparatus of capture to Zuboff's (2019) notions of "surveillance capitalism" and "behavioural surplus". We prefer Deleuze and Guattari's concept of machinic surplus value, which pushes the anthropocentric agenda much further than Zuboff's critique. To a large extent, Zuboff's critique of surveillance capitalism as the exploitation of behavioural surplus still presupposes the liberal idea of non-mediated form of human agency.

[16] Hörl (2013, 129) speaks of a "non-modern mapping of environmentality", "purely relational systems that appear to act as alternative cartographies for a non-modern reframing of our present and future technological world". These non-modern cartographies work as "metamodels for the urgent cosmo-technological reconceptualization of participation as constitutive relationality and therefore too of agency, relationship and relatedness, experience and subjectivity, all of which we need if we are to understand our no-longer-rejected originary environmental condition in a non-reductionist way" (Hörl 2013, 129).

Hui (2017, 2) contends that technology "is not anthropologically universal; it is enabled and constrained by particular cosmologies, which go beyond mere functionality or utility". This means that we should not think of technology as a single, universal, and homogeneous force, but rather as a field inhabited by "multiple cosmotechnics" (Hui, 2017, 2).[17] Citing Philippe Descola, Hui argues that modernity is characterised by an "opposition between culture and nature, and the former's mastery over the latter" (2017, 5). Modern Western thought opposes nature to culture and conceives the former "as a universal ground that is common to all particular cultures" (Lemmens, 2020, 3). An appeal to other ontologies, the so called "ontological turn", is an appeal to the overcoming of this anthropocentric conception of nature. The problem, Hui argues, is that "the question of technics is not sufficiently addressed in the ontological turn" (2017, 6). In most cases, authors place technology on the side of modern Western ontology, and thus close off the possibility of a non-modern and non-Western interpretation of technology beyond the culture-nature divide (Hui, 2017, 6).[18] If the "ontological turn" was aimed at overcoming the nature-culture divide, Hui's concept of cosmotechnics "is designed to overcome modernity's opposition between nature and technology" (Lemmens, 2020, 4). This should not be understood as a defence of tradition or non-technical worldviews motivated by a desire to rehabilitate the past, but rather a project that "explicitly looks at the future and aims to be an imaginative and inventive discipline in search for new cosmotechnics [that is] a plurality of cosmotechnics for the age of the Anthropocene" (Lemmens, 2020, 4; see also Hui, 2020, 64).

As mentioned above, Hörl credits cybernetics with the blurring of the conceptual distinction between nature, culture and technics. At the same time, however, cybernetics "remains a thinking of totalisation" (Hui, 2020, 63). Comsmotechnics, on the other hand, is not concerned with totality (and universalism) but with "technodiversisity"; as such, it tries to think "beyond the totalising effect of cybernetics" (Hui, 2020, 63). It introduces the "question of locality" into the totalising thinking of cybernetics (Hui, 2020, 63). Just as "nature" requires biodiversity, the new planetary ensemble of nature-culture-technics, the "general ecology", requires "technodiversity" (Hui, 2020, 63). The disappearing of a species (biodiversity) is equivalent to the disappearance of different cosmotechnics and the imposition of a universal view of technology. The concept of cosmopolitics is hence presented as a way of refusing this path, challenging the "homogeneous technological future that is presented to us as the only option" (Hui, 2017, 9).

Conclusion: Towards a posthumanist critique of large language models

We argue that the theoretical framework grounded in the above four concepts provides a productive lens for critiquing LLMs on posthumanist terms. Such posthumanist critique serves a twofold function. On the one hand, it enables a departure from current approaches like that of Bender, which—on account of our allegedly exclusive capacities for language, intelligence, and meaning-making—promote human exceptionalism and advocate for a politics of technological governance that seeks to safeguard our humanity and prevent a dehumanizing conflation with machines. On the other, it responds to Hörl's call for a new critique, one that will be capable of better understanding our "digital condition". Each of the four proposed concepts thus addresses an aspect

[17] Hui's concept of cosmotechnics stems from a critical dialogue with both Kant and Stengers. He challenges Kant's "pursuit of the universal", calling for "a certain relativism as the condition of possibility for coexistence" (2017, 2). At the same time, Hui complements Stenger's work on "cosmopolitics" by emphasising the centrality of technology for any "politics to come" (2017, 2).

[18] Similarly, Lemmens (2020, 4) argues that the ontological turn is "an attempt to reconceptualize the relation between the human and the non-human, and hence to go beyond the nature-culture dichotomy that restricts all visions to a parochial Western worldview". Through colonization and modernization, most non-Western cultures have been "enframed by Western technology" in such a way that "the global technological condition has become their destiny" (Lemmens, 2020, 4).

of LLMs from a perspective that both avoids falling back on an anthropocentric standpoint while also accounting for the new mechanisms of power under which these technologies emerge.

Firstly, Hörl's concept of general ecology allows us to consider LLMs as systems of relations that challenge the distinction between culture, nature, and technology. From this perspective, humanity is no longer seen as an agential force that stands above the inert realm of nature and uses technology to transform it and exploit it. Instead, the interconnected multiplicity of human and nonhuman components that feed into the functioning of a language model co-exist on a horizontal ontological plane. These components do not exist as separate and self-contained entities but are constituted only through interactions with their environments and a continuous co-shaping of one another. As such, LLMs should be understood as an "assemblage" (Buchanan, 2021; Hayles, 2023; Lindgren, 2023) that operates through complex interaction between different material and symbolic components.[19] These components include human developers, moderators, and users, mathematical models, training datasets, computational infrastructures (GPUs, servers, cables, etc.), specialized knowledges, socio-technical imaginaries, and vast amounts of electric energy and water-cooling systems. In the specialised literature, experts speak of these assemblages as "ecosystems" necessary to develop, maintain, and deploy "foundation models" such as LLMs (Briggs, 2023). All of this is embedded into wider socio-technical systems of material infrastructure, legal institutions, and capitalist markets. If we follow Buchanan (2021, 144), we can speak of LLMs as "control assemblages", that is, assemblages of human and nonhuman components organised in such a way that they provide leverage for the new mechanism of power that underlie digital platforms. Considering LLMs as control assemblages prevents us from seeing technology in purely instrumental terms and nature as a mere exploitable resource. Additionally, it highlights the relation between LLMs and new datafied technologies of power in the context of digitalisation, or what Hörl calls Environmentality.

Secondly, the concept of machinic agency enables us to go beyond the opposition of humans as autonomous agents and language models as a mindless, probabilistic tool. From the perspective of machinic agency, agential forces are a matter of enabling and disabling connections between human and nonhuman components (see Markelj and Celis Bueno, 2023). As such, agency (intelligence, language, meaning-making, etc.) is not seen as the exclusive domain of human subjects, but as arising from the interaction between different nodes in the LLM assemblage described in the previous paragraph. Framing agency in these terms does not amount to simply attributing anthropocentric agency (sentience, consciousness, self-awareness) to AI systems as is done by the transhumanist vision of a technological singularity. The latter reproduces the anthropocentric view of agency as it posits that such singularity is autonomously created by humans and possesses an agency that is superior but akin to ours (Hui, 2017, 12). As machinic agency locates the capacity to act in the affective relation between humans and non-humans, it challenges the anthropocentric notions of authorship, creativity, and intentionality (Markelj and Celis Bueno, 2023), ideas that seem to have been reinforced by some of the current critiques of *ChatGPT* discussed above. Aligned with Stiegler's theory of "originary technicity", our perspective assumes that our interaction with language models is co-constitutive of both the user and the machine through a reciprocal and recursive process (Hui, 2023). In this sense, current critiques of LLMs that simply call for institutional regulation (private or public) tend to assume an instrumental view of technology that obfuscates the machinic agency of each of the different components of the LLM assemblage.

[19] We use the concept of assemblage in the sense outlined by Deleuze and Guattari (1987). For a thorough discussion of the concept, see Buchanan (2021).

Alternative, a posthumanist critique of LLMs assumes a distributed conception of agency that redefines the terms of any possible regulation. From this perspective, any sustainable and long-term intervention needs to be aware of how each component of the entire ecosystem of LLMs enacts its own agency, reshaping the others in a co-constitutive process. This entails regulating not only the technologies themselves, but the social assemblages "capable of producing them and making use of them" (Deleuze, 1995, 180).

Thirdly, the concept of machinic surplus value allows us to focus precisely on these social assemblages and to examine LLMs as a capitalist apparatus of capture. From this perspective, LLMs can be seen as "control assemblages" that seek to exploit and appropriate the "amplifying connections" that arise from the interactions between its human and nonhuman components. By shifting from a Marxist framework to a posthumanist critique grounded in the concept of machinic surplus value it can be shown that exploitation in current LLMs does not simply concern human intentional effort (e.g., that of OpenAI's engineers and outsourced content moderators), but rather affective relations between human and nonhuman elements which precede the said conscious effort. If we assume that LLMs constitute a "control assemblage" in which agency emerges from relations between its human and nonhuman components, then the traditional notions of labour and surplus value (which concern exclusively the domain of conscious human activity) appear insufficient for a critique of the social assemblages that produce these technologies and put them to work. Mobilising the (posthumanist) concept of machinic surplus value allows us to emphasise that a critique of the political economy of LLMs is not simply a matter of analysing the conflict between human labour and capital but is instead grounded in the examination of the mechanisms of capture of a given surplus. Here surplus is not understood in terms of human (abstract) labour time, but as the process of connective amplification that stems from a specific machinic assemblage. As Hörl puts it, a critique of political economy in the current context of cybernetization cannot be restricted to the exploitation of human labour time but must focus on the processes of "exploitation of relationality". In this sense, large language models allow for new relations of amplification that cannot simply be explained in terms of the automation of labour and the reduction of abstract labour time. As an apparatus of capture, LLMs exploit this relationality in ways that exceed the traditional forms of capitalist valorisation.

Finally, the concept of cosmotechnics urges us to redefine our modes of engagement with LLMs. These modes of engagement should be able to go beyond the universal imperatives of productivity and profit that have been dominating modern mobilizations of technology, and that still posit nature as a "standing reserve" (a mere resource to be exploited by human enterprise). In relation to large language models, Hui argues that their development in the last decade has been dictated by "the competition of technological acceleration and the allures of war, technological singularity, and transhumanist dreams" (2017, 9). Yet, he insists that different techno-ecologies are possible. As our engagement with LLMs inevitably shapes our capacities to act, creating more enabling modes of its deployment is paramount. This demands imagining radical new ways of defining technology beyond the universalising, totalising, and homogeneous framework through which technology has been understood in the Western worldview. This is not an easy task, precisely because of the pervasiveness of this conceptualisation of technology. Our understanding of large language models such as *ChatGPT* remains seized by an anthropocentric worldview that (a) sees technology as a means to transform nature in order to satisfy human needs, and (b) measures this process in terms of productivity and profit (human surplus value). Hui's concept of cosmotechnics is a call for disentangling technology from this unitary model and imagine other modes of collaboration between humans and machines. One outlet for the creation of these new, more empowering and

sustainable technological imaginaries is the practice of counter-memory, which emphasises the excavation of "forgotten or marginalised histories" of technics (Sylvia IV, 2021, 145). Following Sylvia IV (2021), we propose that uncovering these minoritarian narratives of technology and acknowledging their cultural specificity can be a significant resource for posthumanist ethics and politics as they can account for more enabling processes of subjectivation.

The posthumanist critique sketched in this article is an attempt to move in the direction of such new collaborations. By conceptualising LLMs as a "control assemblage" in which agency is distributed between human and nonhuman components, and by articulating its exploitative nature in terms of the capitalization of relationality and the capturing of processes of amplification, we aim at disentangling its critique from an anthropocentric framework. In doing so, we are taking the first step towards a different way of imagining these technologies beyond the human imperatives of exploitation, productivity, and profit.

References

Alaimo, S. (2010). *Bodily Natures*. Indiana University Press.

Barad, K. (2007). *Meeting the Universe Halfway*. Duke University Press.

Bell, E. (2023). A fake news frenzy: Why ChatGPT could be disastrous for truth in journalism. *The Guardian*. https://www.theguardian.com/commentisfree/2023/mar/03/fake-news-chatgpt-truth-journalism-disinformation

Bender, E. M. (2022). Human-like programs abuse our empathy – even Google engineers aren't immune. *The Guardian*. https://www.theguardian.com/commentisfree/2022/jun/14/human-like-programs-abuse-our-empathy-even-google-engineers-arent-immune

Bender, E. M. (2024). Resisting dehumanisation in the age of AI. *Current Directions in Psychological Science*, 33(2), 114-120. https://doi.org/10.1177/09637214231217286

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198. https://doi.org/10.18653/v1/2020.acl-main.463

Bennet, J. (2009). *Vibrant Matter: A Political Ecology of Things*. Duke University Press.

Bode, K., & Goodlad, L. M. E. (2023). Data Worlds: An Introduction. *Critical AI*, *1*(1–2). https://doi.org/10.1215/2834703X-10734026

Bogost, I. (2023). ChatGPT Is Dumber Than You Think. *The Atlantic*. https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/

Braidotti, R. (2013). *The Posthuman*. Polity.

Bridle, J. (2023). The Stupidity of AI. *The Guardian*. https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt

Briggs, J. (2023). LLMs are not all you need. *Pinecone*. https://www.pinecone.io/learn/llm-ecosystem/

Buchanan, I. (2021) *Assemblage Theory and Method*. Bloomsbury.

Burgess, M. (2023). ChatGPT Has a Big Privacy Problem. *Wired*. https://www.wired.com/story/italy-ban-chatgpt-privacy-gdpr/

Celis Bueno, C. (2024). Postcapitalist Surplus Value: Humans, Machines, Information. In A. Pratihar (Ed.), *Deleuze, Guattari and the Schizoanalysis of Post-Neoliberalism* (pp. 55–75). Bloomsbury Academic.

Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory, Culture & Society*, *28*(6), 164–181. https://doi.org/10.1177/0263276411424420

Chiang, T. (2023). ChatGPT is a Blurry jpeg of the Web. *The New Yorker*. https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web

Chomsky, N., Roberts, I., & Watumull, J. (2023). Noam Chomsky: The False Promise of ChatGPT. *The New York Times*. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html

Coeckelbergh, M. (2022). *The political philosophy of AI: An introduction*. Polity press.

Coeckelbergh, M., & Gunkel, D. J. (2023). ChatGPT: Deconstructing the debate and moving it forward. *AI & SOCIETY*. https://doi.org/10.1007/s00146-023-01710-4

Crawford, K. (2021). *Atlas of AI*. Yale University Press.

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. *Frontiers in Public Health*, *11*, 1166120. https://doi.org/10.3389/fpubh.2023.1166120

Deleuze, G. (1988). Spinoza: Practical Philosophy, City Lights Books.

Deleuze, G. (1995). *Negotiations*. Columbia University Press.

Deleuze, G., & Guattari, F. (1983). *Anti-Oedipus*. Continuum.

Deleuze, G. & Guattari, F. (1987). *A Thousand Plateaus*. University of Minnesota Press.

Derrida, J. (1976). *Of Grammatology*. Johns Hopkins University Press.

Emsley, R. (2023). ChatGPT: These are not hallucinations – they're fabrications and falsifications. *Schizophrenia*, *9*(1), https://doi.org/10.1038/s41537-023-00379-4

Foucault, M. (2009). *Security, Territory, Population: Lectures at the Collège de France 1977-1978*. Palgrave Macmillan.

Guattari, F. (2014). *The three ecologies*. Bloomsbury Academic.

Gupta, A., Atef, Y., Mills, A., & Bali, M. (2024). *Assistant, Parrot, or Colonizing Loudspeaker? ChatGPT Metaphors for Developing Critical AI Literacies*. https://doi.org/10.48550/ARXIV.2401.08711

Hansen, M. (2009). System-Environment Hybrids. In B. Clarke & M. Hansen (Eds.), *Emergence and Embodiment* (pp. 113–142). Duke University Press.

Hansen, M. B. N. (2015). *Feed-forward: On the future of twenty-first-century media*. University of Chicago Press.

Hao, K. (2019). Training a single AI model can emit as much carbon as five cars in their lifetimes. *MIT Technology Review*. https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/

Hayles, N. K. (2022). Inside the Mind of an AI: Materiality and the Crisis of Representation. *New Literary History*, *54*(1), 635–666. https://doi.org/10.1353/nlh.2022.a898324

Hayles, N. K. (2023). Technosymbiosis: Figuring (Out) Our Relations to AI. In J. Browne, S. Cave, E. Drage, & K. McInerney (Eds.), *Feminist AI* (1st ed., pp. 1–18). Oxford University PressOxford. https://doi.org/10.1093/oso/9780192889898.003.0001

Herbrechter, S. (2018). Critical Posthumanism. In R. Braidotti & M. Hlavajova (Eds.), *Posthuman Glossary* (pp. 94–96). Bloomsbury.

Hörl, E. (2013). A Thousand Ecologies: The Process of Cyberneticization and General Ecology. In D. Diederichsen & A. Franke (Eds.), *The Whole Earth. California and the Disappearance of the Outside* (pp. 121–130). Sternberg Press.

Hörl, E. (2017). Introduction to general ecology: The ecologization of thinking. In *General Ecology* (pp. 1–73). Bloomsbury.

Hörl, E. (2018). General Ecology. In R. Braidotti & M. Hlavajova (Eds.), *Posthuman Glossary* (pp. 172–175). Bloomsbury.

Hörl, E. (2021). Critique of Environmentality: On the World-Wide Axiomatics of Environmentalitarian Time. In E. Hörl, N. Y. Pinkrah, & L. Warnsholdt (Eds.), *Critique and the digital* (pp. 109–146). Diaphanes.

Hörl, E., Pinkrah, N. Y., & Warnsholdt, L. (2021). Introduction. In E. Hörl, N. Y. Pinkrah, & L. Warnsholdt (Eds.), *Critique and the digital* (pp. 7–24). Diaphanes.

Hsu, T., & Thompson, S. A. (2023). Disinformation Researchers Raise Alarms About A.I. Chatbots. *The New York Times*. https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html

Hui, Y. (2017). Cosmotechnics as Cosmopolitics. *E-Flux Journal*, *86*, 1–11.

Hui, Y. (2020). Machine and Ecology. *Angelaki*, *25*(4), 54–66. https://doi.org/10.1080/0969725X.2020.1790835

Hui, Y. (2023). ChatGPT, or the Eschatology of Machines. *E-Flux*, *137*. https://www.e-flux.com/ journal/137/544816/chatgpt-or-the-eschatology-of-machines/

Kanungo, A. (2023). The Green Dilemma: Can AI Fulfil Its Potential Without Harming the Environment? *Earth.Org*. https://earth.org/the-green-dilemma-can-ai-fulfil-its-potential-without-harming-the-environment/

Lazzarato, M. (2014). *Signs and Machines: Capitalism and the production of subjectivity*. Semiotext(e).

Lemmens, P. (2020). Cosmotechnics and the Ontological Turn in the Age of the Anthropocene. *Angelaki*, *25*(4), 3–8. https://doi.org/10.1080/0969725X.2020.1790830

Ligozat, A.-L., Lefevre, J., Bugeau, A., & Combaz, J. (2022). Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions. *Sustainability*, *14*(9), 5172. https://doi.org/10.3390/su14095172

Lindgren, S. (2023). *Critical theory of AI*. Polity Press.

Marchand, J. S. (2018). Non-human Agency. In R. Braidotti & M. Hlavajova (Eds.), *Posthuman Glossary* (pp. 292–295). Bloomsbury.

Markelj, J. & Celis Bueno, C. (2023) Machinic agency and datafication: Labour and value after anthropocentrism. *Convergence, 30(3),* 1058–1075. https://doi.org/10.1177/13548565231166534

Marx, K. (1976). *Capital: A Critique of Political Economy* (Vol. 1). Harmondsworth; Penguin; New Left Review.

Metz, C. (2023). Chatbots May 'Hallucinate' More Often Than Many Realize. *The New York Times*. https://www.nytimes.com/2023/11/06/technology/chatbots-hallucination-rates.html

Motoki, F., Pinho Neto, V., & Rodrigues, V. (2024). More human than human: Measuring ChatGPT political bias. *Public Choice*, *198*(1–2), 3–23. https://doi.org/10.1007/s11127-023-01097-2

Pasquinelli, M. (2015). Italian Operaismo and the Information Machine. *Theory, Culture & Society*, *32*, 49–68.

Perrigo, B. (2023). Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic. *Time*. https://time.com/6247678/openai-chatgpt-kenya-workers/

Perucica, N., & Andjelkovic, K. (2022). Is the future of AI sustainable? A case study of the European Union. *Transforming Government: People, Process and Policy*, *16*(3), 347–358. https://doi.org/10.1108/TG-06-2021-0106

Plato. (1997). *Phaedrus*. Cambridge University Press.

Rammert, W. (2008). Where the Action is: Distributed Agency between Humans, Machines, and Programs. In U. Seifert, J. H. Kim, & A. Moore (Eds.), *Paradoxes of Interactivity* (pp. 62–91). Transaction.

Rees, T. (2022). Non-Human Words: On GPT-3 as a Philosophical Laboratory. *Daedalus*, *151*(2), 168–182. https://doi.org/10.1162/daed_a_01908

Rouvroy, A., & Berns, T. (2013). Algorithmic Governmentality and the Prospects of Emancipation. *Reseaux*, *177*, 163–196.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, *3*, 417–457.

Stiegler, B. (1998). *Technics and Time, 1: The Fault of Epimetheus*. Stanford University Press.

Sylvia IV, J.J. (2021). 'Posthuman Media Studies', in Journal of Posthumanism, December 2021, Vol.1, No.2, pp. 139-151. https://doi.org/10.33182/jp.v1i2.1360

Tufecki, Z. (2022). What Would Plato Say About ChatGPT? *The New York Times*. https://www.nytimes.com/2022/12/15/opinion/chatgpt-education-ai-technology.html

Valdivia, A. (2022). Silicon Valley and the Environmental Costs of AI. *Political Economy Research Centre*, *8*. https://www.perc.org.uk/project_posts/silicon-valley-and-the-environmental-costs-of-ai/

Weatherby, L. and Justie, B. (2022). Indexical AI. *Critical Inquiry*, 48(2), pp.381-415.

Weil, E. (2023). You Are Not a Parrot. *New York Magazine*. https://nymag.com/intelligencer/article/ai-artificial-intelligence-chatbots-emily-m-bender.html

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs.