AI is a Trojan Horse – A Rationale 🐎

Dominik Schlienger
Department for Music and Technology
University of the Arts Helsinki
Helsinki, Finland
0009-0002-4340-9116

***Abstract*—Mainstream media report that governments and business leaders see AI as an "extinction-level threat" to humans. In posthumanism, we find, besides conceptional tools for thinking technology in an indeterministic way, also argumentations for *singularity*, *accelerationism* and appeasement towards *strong AI*. Notwithstanding the heterogeneity of the field, the eminent posthumanist claim that autonomous machines can be a conscient *other* is a fallacy. An analysis of computer language as Derridean writing shows that computational, supposedly autonomous machines run on algorithmic, linguistic, *written* code. As such, they are extensions of human cognition. To proclaim machines conscient and autonomous, is hence not just misleading per se, but disguises the human agency that uses the AI-autonomy as a proxy.**

***Keywords*—Code as Derridean writing, Fallacy of AI autonomy, AI ethics. Posthumanism and AI**

I. Introduction

The Mainstream media report that governments and business leaders see AI as an "extinction-level threat" to humans [1]–[3]. What is more, a 'founder figure' who resigned from a senior position with a leading player in AI named the "growing dangers from developments in the field" as a reason for his departure [4]. Through this and similar other stories, the media discourse gained credibility; further supported by governments reacting to the perceived AI threat by turning to 'experts' for advice on policy [5].

In computer science, the general discourse of AI as a threat to humanity is reflected in an increased interest in ethics in technology, as the call for this paper testifies. Along with the growing understanding for the ethical question in technology development, comes a desire to understand where we are going with these *emerging* technologies. The prognostics range from cautious to the grotesque:

In [6] Timnit Gebru, Emily M. Bender, *et al.* give an articulate insider view on how we should proceed with particular technologies, namely large language models referred to *elsewhere* as "artificial intelligence" — a term that does not occur in the paper. Rather than painting the devil on the wall, they call for considerate, ethical action, and for awareness of the environmental impacts of the energy-intensive technology.

In stark contrast to that specific attitude, the authors of [7] make sweeping assumptions on the future of "AI" claiming credibility by stealth of numbers, surveying 2,778 researchers who had published in "top-tier artificial intelligence venues". But no definition of the term AI itself is given. In the questionnaires, the researchers were asked about the feasibility of alleged forms of AI, in the immediate and near future, ranging from "High-Level Machine Intelligence", to "Full Automation of Labor", notwithstanding the extremely hypothetical nature of both concepts given the state of the art. Interestingly, no data is provided as to the relation of the subjects' specialisation in "AI" and the specific answers given: Arguably, a credible expert in, say, applied statistics could give their fully speculative opinion on mechanical engineering tasks. A closer look at the data illustration also shows that the predictions diverge so extremely, that no individual subject represents the estimated average even closely. That did not keep the authors from discussing the outcomes as indicative data.

Incredulously but in a similar vein, the authors of [8] claim that "AI" may be "capable of generating an ethic to save the planet" as "there appears a new, virtual environment, capable of making complex analyses and replacing the human element in making decisions".

The mainstream view in the media and its reflection in a segment of computer science, seems to be underpinned by a third strand of the AI discourse — a literal reading of posthumanist concepts. Posthumanism is instrumental in much of the academic writing on AI and ethics. In posthumanism we find, besides conceptional tools for thinking technology in an indeterministic way [9]–[11], also argumentations for *singularity* [12], *accelerationism* [13] and appeasement towards *strong AI* [14].

What is intriguing is that calls for restraint in developing AI come from all corners of society. Except for proponents of *accellerationism* [13] who think that all problems occurring in AI development can be solved by more development, there is an uncanny consenus on the need for regulation on AI, and, concurrently a fatalist compliance with the inevitability of *singularity*, the idea that soon AI will be more powerful than humans, and eventually, take control of the world. So what is going on?

In what follows, I aim to deconstruct the criticism on AI to show that its constituent parts are rooted in two mutually exclusive rationales. Separated by a fine dividing line, we can see how criticism on AI criticises *how we do AI* on the one side, and *what AI does to us* on the other. I will argue, that

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

the latter form of criticism is based on a misconception of technology. What is more, I will show how the posthumanist argument for technology as a *conscient other* is based on the logical fallacy that code could write itself.

Although I do not concur with the type of criticism of posthumanism that aims to defend the Enlightenment at best [15], and white male hegemony at worst, I still attempt to draw attention to what I see as a weakness in the argumentation underpinning the current discourse according to which computational machines could be autonomous to the extent the media portrays. In contrast, I assert that the reliance of computational machines to run on algorithmic code makes the concept of singularity logically impossible: Through language, machines are extensions of the cognitive practices that constitute the language they run on [16] The computer is to the brain what the hammer is to the hand.

II. A short and Incomplete Overview of Posthumanism

Within the scope of this paper, a thorough overview of posthumanism shall not be attempted. But I dare to roughly categorise two types of posthumanists. Concurring with the argumentation in [10], Bruno Latour, Michel Foucault, Karen Barrad, and Donna Haraway, form a first group of thinkers for whom posthumanism expresses a critique on humanist *exceptionalism*. For that group, posthumanism stands for a strand of thinking whereby "we have never been only human".

In contrast, I see a second group with Nick Bostrom [17], Nick Land [18], and Benjamin H. Bratton [19], and others. But I concentrate here on N. Katherine Hayles's work, where the "human tradition is necessary to allow the emergence of the post-human through a conceptualisation of intelligence as co-produced with intelligent machines" [20]. There are no essential differences or absolute demarcations between bodily existence and computer simulation. This "privileges information over materiality, considers consciousness as an epiphenomenon and imagines the body as a prosthesis for the mind"[ibid.]. Hayles states that "cognitive technologies are now a potent force in our planetary cognitive ecology" and that there are "rapidly escalating complexities created by the interpenetration of cognitive technologies with human systems" The two last tenets, for Hayles "are not debatable*[sic]*" [14]. Further, she asserts as given that there are technical autonomous systems; cognition in technical systems; and limited human decision power in such systems. These positions are not uncommon in posthumanist thought, as recent, and more extensive overviews confirm, some of them acquiescing to the logical consequences of such thinking [12], [21], some exercising it as a conceptual possibility [13].

In summary, there are two groups of posthumanists discernible, one of which thinks the post-human as a means to conceptualise the *anthropos* as relational, non-exceptional, *holobiontic* [22] entity amongst others, and a second group, which conceptualises the post-human as a transformation from human to *something post-human* via *technology*. In what follows, I will assert, by example of Hayles' argumentation, that autonomous cognition in machines, far from being "not debatable", is a *non sequitur*.

III. Methodology

I am using media coverage of AI on the one hand, and on the other the notion of "posthumanism" as a recent development in philosophy, media study and science and technology studies as my sources, including recent works that position posthumanism in response to the latest developments in AI. In continuation from my thesis on agency in technology [16], I contrast the claim of cognitive autonomy in both media coverage and posthumanism with my argument that through the reliance on, and presence of *code* in all present — and arguably future computational machines, actions that could be perceived as cognitively autonomous are an extension of human meaning making. My argumentation applies a post-structuralist linguistic comparison of natural language with computer language based on Umberto Eco's Theory of Semiotics [23], and Jacques Derrida's On Gramatology [24]. In my thesis I developed my argumentation on the basis of my practical experience in technology development. Here I compare the outcome of the thesis, that technology is a socio-material, performative practise, to the current discourse on AI, in the media and recent philosophical developments.

To delineate the methodology more explicitly: I apply a pragmatic and analytic philosophical approach to concepts of linguistics and semiotics as found and applied in the selection of media reports, academic literature on AI and ethics, and philosophical discussions that constitute my data. My pragmatism originates from a practical background in technology development, and an adoption of analytical principles from anthropology as applied in science and technology studies and actor network theory.

IV. Code as Derridean Writing

Let us start with the working hypothesis that, if — as I believe, AI is a tool like any other technical tool, and as such, not a cognising *other* but an extension of *our* embodied cognition, then the dangers of AI lie with *what we do with AI*, and *singularity* is consequently not possible, as technical agency in the machines that we deploy is firmly and causally linked to *our* actions. On the contrary, if it can be logically argued that the cognition in AI is not human, then it is possible that AI *does things to us*. In support of the working hypothesis, I further propose that through *code*, through computer languages, and algorithms written in such languages, computational machines, effectively, are linguistic devices and hence have grammar, syntax, and semantics just like natural languages. They are hence a system of human meaning making.

In [25] N. Katherine Hayles argues that computer languages are *not* languages in the sense of natural languages, which consequently makes it possible to argue that the meaning making in computational machines is effectively disconnected from human meaning making. Yet, as I discuss elsewhere in more detail [16], a linguistic analysis shows, that computer

languages indeed do constitute languages, and in particular, constitute Derridean *writing*, based on Derrida's work in *On Grammatology* [24]. Here a succinct summary of my argumentation:

Derrida adopts from Saussure the binary oppositions that all speech or text has to articulate if it intends to make sense [26]. Herein lies a conceptual analogy to binary code: on the technical (procedural) level, all code works in a similar way, and the complexity of sense-making can be explained through the proliferation of multiple simple structures. The complexity or limitation of meaning lies not in the binary opposition in sense-making, but in the complexity of the available signal-paths.

That language itself is not dissimilar to a computational machine can be seen from Umberto Eco's *Watergate Model* as proposed in his Theory of Semiotics [23]. The model consists of signal paths that link a watergate, a *conveying system*, with a control station downstream, the *conveyed system*. The code is what connects them, element by element. The elements of the former become the expression of the latter, and the latter becomes the content of the former. For Eco, there are no signs, only sign-functions. This defines the difference between a signal and a sign: "A signal is a pertinent unit of a system that may be an expression system ordered to a content but could also be a physical system without any semiotic purpose" A signal can be entirely free of meaning. But when the signal is recognised as an antecedent of a foreseen consequent it may be viewed as a sign.

Further, computational machines and their programs can be understood as a trace of their purpose. For Derrida, all traces are the consequences of *writing*. Conversely, writing leaves traces. What is written, remains, and can be transported, read later, spoken, activated, compiled, executed, debugged, updated, or left to (code-)rot. One essentially Derridean aspect of this type of universal writing comes from the notion of différance, an intentional misspelling of 'différence', which Derrida uses to indicate firstly, that by writing the word differently (as a grapheme) there is an additional meaning to the conventional understanding of the word when we hear it as a phoneme (they are pronounced exactly the same way). Secondly, he means to indicate the double meaning of the French word 'différer' which translates to differ as well as to defer. Différance in the sense of to differ stands for the binary sense-making encapsulated in Saussure's linguistic sign. Différance stands for defer, however, to indicate that meaning can be deferred through chains of signifiers, and as a grapheme, in particular, also postponed.

Therefore, computer language and written natural language do not conceptually differ. Furthermore, due to Iterability, is is possible for a written text to remain readable after the person who wrote it is gone [18]. It can also be taken out of its original context and, when reiterated, work in a different context. The writing itself remains, it is a trace. like a machine at rest. That a sentence means something entirely different according to whom to, when, and how I say it does not change the literal text, the code. Further, Derrida insists that "there is nothing outside context", so the signifier's *conditio sine qua non* is context.

To recap my core points, Computational machines are linguistic systems, congruent with Eco's Watergate-Model. The code they run on is iterable in a Derridean sense, and it can be deferred to another place and time.

V. Il n'y a pas de Hors-Texte

However, using the same sources (Derrida and Saussure), Hayles concludes that computer languages are *not* languages in a Derridean sense, based on the argument that iterability can *not* be applied to code, "where contexts are precisely determined by the level and nature of the code. [...] Code may be rendered unintelligible if transported into a different context—for example, into a different programming language or a different syntactic structure within the same language"

So how does Hayles and my reading of Derrida and Saussure differ? I believe the difference becomes visible if we recall Eco's watergate as a model for decision making — we can see that the system constitutes the code, in form of a consequent and a precondition, a conveyor of (meaningful) decision making. But the system (the code) is not the meaning itself. For Hayles, if "contexts are precisely determined by the level and nature of the code" the meaning *is* the code. This is not congruent with Derrida's and Saussure's linguistic model even if she bases her argumentation nominally on them.

Here, in a nutshell, lies the fallacy of the argumentation applied to make plausible the autonomously cognising computational machine: It is the assumption, that in computer languages, in contrast to any other languages, meaning could be hard-coded into the signifier. But this is not the case: The whole machine, including its "dependencies", like operating system, hardware, and firmware constitute the code (the signal-paths of the watergate), and hence the whole machine is part of the iterable writing, part of the signifier, the trace.

Hayles implies that for code to constitute writing in a Derridean sense, the code would need to be understood by any system. But this contradicts Derrida by implying that code, as the signifier, should *mean* the same (rather than simply *be* the same) in every iteration: The signified would need to stay constant while the signifier changed through iteration. However, this is neither true for code (it wouldn't compile) nor natural language; a text in French written 200 years ago means something slightly different if read today. What is more, its meaning might be opaque but readable for Italians, but unintelligible to most English speaking people. But we would not expect it to make sense as an English text. English, as a language, is an essential component of the constituting code that is used for English texts, like a specific compiler will only compile code written for it.

To put the analogy to the test: Whatever a machine outputs successfully, be it the words "Hello World" or a material 3D print, if it assembles a car, hits a nail, or cuts a lawn — its meaning responds to, *conveys* the intention why the program was written in the first place. For both, machine and language, context is everything. This is particularly evident when looking

at language-as-machines *in extremis*, that is, chatbots based on *large language models*: As computer scientist and AI expert Timnit Gebru and linguist Emily M. Bender quite clearly state [6], the chatbot, as a "stochastic parrot", has no idea of what it "says", and they warn: "The tendency of human interlocutors to impute meaning where there is none can mislead both researchers and the general public into taking synthetic text as meaningful."

The decisions of a technical system are the encoded decisions of the human cognition behind it; Hayles' "cognitive abilities of complex technical systems interpenetrating human complex systems" are not dangerous *per se*, but a tautology in logical terms: Only through a muddling of signifier and signified is it possible to argue that the meaning making in computational machines is in any way disconnected from the meaning making of the human agent building, deploying, and activating that machine. What is more, the fact that machines, just like language and any other type of code, need performing — relational activating, to make meaning, makes the point even more salient that *we do AI* in a certain way.

The (directional) process from an antecedent to a consequent already allured to in the description of the watergate-model is instrumental in the meaning making. Let us call this, rather inelegantly, run-time; for meaning to be constituted, there need to be an antecedent and a consequent activated through code which *runs*.

VI. Cognition at Run-time

The fact that code does not run itself, that computational machines need to be powered up, that they only work at *run-time* means that, if machine did something which could equate to cognition, they would only do this while they *run*. So, one could argue that run-time as *motor action* is a precondition for cognition [27]. Reversely, it means that every entity that *runs* could in theory, potentially cognise. On this basis, motor-neural activity is enough for basic cognitive activity. To contest this would be an anthropocentric view, it would be humanist exceptionalism.

Interestingly there would be no problem *per se* if machines had cognitive abilities; we could interact with them, more or less aptly, like we do with other cognitive beings we encounter. To some extent Donna Haraway's Cyborgs exemplify such encounters. In her epochal essay "A Cyborg Manifesto" [28], she makes the case for embracing a shared kinship with machines. But lo and behold, Donna Haraway writes in the introduction to the "Haraway Reader" [9] in reference to Bruno Latour [29] that "we have never been human, much less man. [...] Nonetheless, in my view, people are human in at least one important sense. We are members of a biological species, Homo sapiens. That puts us solidly inside science, history, and nature". The kinship to machines in Harway's work, can thus be read as manifestations of machines-as-human figurations, rather than humanity being devolved by machines-as-non-human *conscient other*.

The performative nature of technology is also paramount in anthropologist Lucy Suchman's work on autonomous machines [30] which she delineates as human-machine *re-configurations* leaning on Haraway's *figurations*. Suchman conceptualises technologies as socio-material arrangements, an essential term to understand technologies as socially constituted, materially negotiated actions. What is more, she shows how in automatons human action is never replaced, just displaced — preempting a description of the chatbots of today in foresight, where the vast data of the internet is concatenated in surprising ways along algorithmic rules established as technological, and more or less ethical practices.

VII. The Level Playingfield

That what makes an object *technical* is not the object itself, is essential for French philosopher and anthropologist Bruno Latour's description of the "technical mode of existence" [31] epitomised in his realisation that "aeroplanes do not fly, but airlines do" [32].

Yet, Latour's emphasis on non-human actants as equal actors in the level playing field of the actor network [ibid.] is often used as an argument for the conceptual possibility of singularity. I vehemently oppose that idea for the following reason: The level playing field of the actor network includes all actions, also human action. And even if we devise humans as holobionts, an honest analysis of the Anthropocene shows that our (technical) actions leave and left distinctive and material traces. These traces, the materiality of computational machines, the servers, the processors, and the rare earths dug out to build them, are a Derridean *grapheme*. If we try to think singularity in this context, it is hard to see how historically, materially and logically, it could be anything else than an expression of human meaning making.

For Gebru and Bender the realisation that no understanding takes place in AI chat-bots is key: "languages are systems of signs, i.e. pairings of form and meaning. But the training data for Language Models is only form; they do not have access to meaning." [6]. What is valid here for the specific case of chatbots, makes my general point here quite clear too, if applied to the level of code running computational machines in general.

VIII. Concluding Remarks

Returning to the two types of criticism on AI, I tried to make the case for the type of criticism that questions *how we do AI*. And I showed that criticism of *what AI does to us* misses the point: It turns AI into a Trojan horse for a (human) proxy agent by proclaiming AI autonomous. But who is to gain from such an endeavour?

Intriguingly, the criticism of the second type comes from a corner of society that one would expect to have a vested interest in the success of AI. Why would those companies that have pushed the development of "AI" to the point where they now allege that we will soon achieve singularity now warn of it? I'm not on my own with this question, Gebru is quoted as having made the very same point at the RightsCon conference in 2023: "Ascribing agency to a tool is a mistake, and that is a diversion tactic. And if you see who talks like

that, it's literally the same people who have poured billions of dollars into these companies." [33]

There is indeed a plausible motive for why such companies might want to be a hidden proxy agent in AI, as *everything AI* can be marketed as a potent power: A power harnessed by nobody but the companies themselves, as proprietors of the closed-source software. But the power of autonomous AI only exists as long as we believe in it. For that purpose, an official warning about AI's potency can only help. Yet, at the risk of a badly mixed metaphor, the Trojan horse might be wearing the emperor's new clothes here, and it is time to call the bluff.

Gebru, Bender et al's. paper [6] stresses that large language model chat-bots favour the existing hegemonic viewpoint, at the cost of diversity and social justice. Saying so as an employee of the same big player in AI mentioned in the beginning of this paper, is presumed to be the reason why Timnit Gebru lost her job as technical co-lead of the Ethical Artificial Intelligence Team of said company.

To call the bluff, we need to re-own the agency in technology and rethink the way we develop AI. This insight resonates with Hannah Arendt's ethical imperative to "think what we're doing", to not rely on "artificial machines to do our thinking and speaking". Failing that, we are "thoughtless creatures at the mercy of every gadget which is technically possible, no matter how murderous it is" [34]. That the agency in AI is not autonomous but an extension of human cognition, provides a trojan horse for whoever is controlling the machine.

I am neither disputing the potential dangers of technologies, nor questioning their beneficial possibilities. But I contest the deterministic conception of technology that attributes the dangers to the technology itself. The fear of AI relies on the construct of autonomous machines as a malevolent other. As machines are socio-material arrangements [30], the harms and dangers of AI are the direct result of (sloppy) human thinking. To attribute them to an external, *deus ex machina* may be convenient for the hegemony that profits from that fear, but it is a cop-out from collective responsibility. The computer is to the brain what the hammer is to the hand. To exploit posthumanism for a philosophical *placet* for the endeavour to push responsibility for our technologies away from individuals and society just shows how far into society the AI Trojan horse has advanced.

References

[1] N. Y. Times, "AI poses 'risk of extinction,' industry leaders warn," 2023. [Online]. Available: https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html

[2] M. Egan, "AI could pose 'extinction-level' threat to humans and the us must intervene, state dept.-commissioned report warns," 2024. [Online]. Available: https://edition.cnn.com/2024/03/12/business/artificial-intelligence-ai-report-extinction/index.html

[3] C. Vallance, "Powerful artificial intelligence ban possible, government adviser warns," 2023. [Online]. Available: https://www.bbc.com/news/technology-65779181

[4] B. Z. K. . C. Vallance, "Ai 'godfather' geoffrey hinton warns of dangers as he quits google," 2023. [Online]. Available: https://www.bbc.com/news/world-us-canada-65452940

[5] N. Omer, "Sunak, musk and ai: what we learned from the bletchley park summit," 2023. [Online]. Available: https://www.theguardian.com/technology/2023/nov/03/rishi-sunak-elon-musk-ai-summit-what-we-learned

[6] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?" in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 610–623. [Online]. Available: https://doi.org/10.1145/3442188.3445922

[7] e. a. Katja Grace, Harlan Stewart, "Thousands of ai authors on the future of ai," 2024.

[8] L. D. Milici and M. R. Milici, "Is ai capable of generating an ethic to save the planet and contemporary society?" *Proceedings*, vol. 81, no. 1, 2022. [Online]. Available: https://www.mdpi.com/2504-3900/81/1/87

[9] D. J. Haraway, *The Haraway Reader*. New York Routledge, 2004.

[10] O. Pyyhtinen and S. Tamminen, "We have never been only human: Foucault and latour on the question of the anthropos," *Anthropological Theory*, vol. 11, no. 2, pp. 135–152, 2011. [Online]. Available: https://doi.org/10.1177/1463499611407398

[11] K. Barad, *Meeting the Universe Halfway*. New York, USA: Duke University Press, 2006. [Online]. Available: https://doi.org/10.1515/9780822388128

[12] R. Nath and R. Manna, "From posthumanism to ethics of artificial intelligence," *AI Soc.*, vol. 38, no. 1, p. 185–196, sep 2021. [Online]. Available: https://doi.org/10.1007/s00146-021-01274-1

[13] J. Brusseau, "Mapping ai avant-gardes in time: posthumanism, transhumanism, genhumanism," *Discover Artificial Intelligence*, vol. 3, no. 1, p. 32, Oct 2023. [Online]. Available: https://doi.org/10.1007/s44163-023-00080-6

[14] N. K. Hayles, *Unthought: The Power of the Cognitive Nonconscious*. University of Chicago Press, 2017.

[15] T. Osborne and N. Rose, "Against posthumanism: Notes towards an ethopolitics of personhood," *Theory, Culture & Society*, vol. 41, no. 1, pp. 3–21, 2024. [Online]. Available: https://doi.org/10.1177/02632764231178472

[16] D. Schlienger, *Developing Alps: Notes on Agency in Technology*. University of the Arts Helsinki, 2023. [Online]. Available: https://urn.fi/URN:ISBN:978-952-329-287-1

[17] N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press, 2014.

[18] K. Ansell-Pearson, J. McCahery, and N. Land, *Machinic Postmodernism*, ser. Published in Association with Theory, Culture and Society Series. SAGE Publications, Limited, 1996. [Online]. Available: https://books.google.fi/books?id=GBTvkQEACAAJ

[19] B. H. Bratton, "Planetary sapience," 2024. [Online]. Available: https://antikythera.org/planetary-sapience

[20] N. K. Hayles, *How We Became Posthuman:Virtual Bodies in Cybernetics, Literature and Informatics*. Chicago: University of Chicago Press, 1999.

[21] D. Tüfekci Can, "Reinterpreting human in the digital age: From anthropocentricism to posthumanism and transhumanism," *Journal of Educational Technology and Online Learning*, vol. 6, no. 4, p. 981–990, 2023.

[22] S. Keto and R. Foster, "Ecosocialization – an ecological turn in the process of socialization," *International Studies in Sociology of Education*, vol. 30, no. 1-2, pp. 34–52, 2021.

[23] U. Eco, *A Theory of Semiotics*, ser. A Midland Book. Indiana University Press, 1979.

[24] J. Derrida and G. Spivak, *Of Grammatology*, ser. Of Grammatology. Johns Hopkins University Press, 2013.

[25] N. K. Hayles, *My Mother Was a Computer: Digital Subjects and Literary Texts*. University of Chicago Press, 2010.

[26] F. de Saussure, W. Baskin, P. Meisel, and H. Saussy, *Course in General Linguistics*. Columbia University Press, 2011.

[27] M. Jeannerod, *Motor Cognition: What Actions Tell the Self*, ser. Motor Cognition: What Actions Tell to the Self. OUP Oxford, 2006.

[28] D. J. Haraway, *A Cyborg Manifesto : Science, Technology, and Socialist-feminism in the Late Twentieth Century*. Socialist Review (US), 1985. [Online]. Available: https://books.google.fi/books?id=M41CcgAACAAJ

[29] B. Latour, C. Porter, and H. U. Press, *We Have Never Been Modern*. Harvard University Press, 1993.

[30] L. Suchman and L. A. Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions*, ser. Learning in Doing: Social, Cognitive and Computational Perspectives. Cambridge University Press, 2007.
[31] B. Latour, *An inquiry into modes of existence : an anthropology of the moderns / Bruno Latour ; translated by Catherine Porter*. Harvard University Press Cambridge, Massachusetts, 2013.
[32] ——, *Reassembling the Social:An Introduction to Actor-Network-Theory: An Introduction to Actor-Network-Theory*, ser. Clarendon Lectures in Management Studies. OUP Oxford, 2005.
[33] T. Ryan-Mosley, "It's time to talk about the real ai risks," 2023. [Online]. Available: https://www.technologyreview.com/2023/06/12/1074449/real-ai-risks/
[34] H. Arendt, *The Human Condition*, ser. Anchor Books. University of Chicago Press, 1958.