

# Three limitations of Algorithmic Reason - Steering the human mind in the twenty first century.

Paulo Castro

Centre for the Philosophy of Sciences of the university of Lisbon (CFCUL), Lisbon, Portugal

[paulo.castro.pi@gmail.com](mailto:paulo.castro.pi@gmail.com)

**Abstract:** Artificial Intelligence has pervaded contemporary societies in almost every way as an externalized, fragmented, and optimized form of rationality. I call this externalized mode of thought “algorithmic rationality” and to the ideology favoring it, “algorithmic reason”. Although algorithmic reason original goal was to facilitate the arising of a highly participative and global Culture, fitting all citizens in a dynamic democratic society, History has it that AI technology would be ceased by consumer logic and computational propaganda. I discuss the economical, epistemological, and political implementation of algorithmic reason, introducing three cases. I argue that such implementation comprises a cybernetic loop, involving a centralized AI and its instrumentalized users. Commenting upon the ubiquity of such loop, I introduce three limitations of algorithmic reason. The first two are of a computational nature. The third owes its presence to a cybernetic loop, producing a steering effect on the human mind and promoting a cultural flattening effect. This, I conclude, may result in the impoverish of creativity, critical thought, and intellectual curiosity.

**Keywords:** Artificial Intelligence, Algorithmic Reason, cultural steering, human mind steering, cybernetic loop, computational propaganda, AI Ethics.

## 1. The raise of algorithmic rationality.

Human reason in its classical expression, the one taught in classrooms for about two centuries, emerged in ancient Greece as a natural tool to provide logical and insightful arguments for political activity. Much later, reason was again used as a diplomacy tool to gather differing religion parties at the same table, soon after the Thirty Years' War, providing a way for religious tolerance and peace a maintenance (Buhr, 2004). Of course, one should also celebrate the efforts of Galileo, Kepler, Copernicus, and many other emancipated spirits, nurturing the scientific revolution, at the time, a major step in human reason evolution.

But now, we are confronted with still another evolving moment in the history of rational thought. In this paper I will speak of some of its previewed undesirable consequences. I will begin my exposition by invoking three main aspects of what I call the rise of “algorithmic rationality”, which I consider to be one of the major events in recent History. It requires first, a necessary reference to the cyberculture movement in the US, when the personal computer was credited as a contestation symbol against the establishment, both in its political and economic dimensions. Next, I will refer to that same movement’s failure, as it ended being ceased by an emerging techno-capitalism, recently called “surveilling capitalism”. Finally, and paving the way to further topics, I will briefly introduce the History of the “resold human attention” as a critical technique for perhaps the most stringent application of algorithmic rationality in contemporary civilization.

The raise of an algorithmic rationality, in the mid XX Century, initially aimed to perform an ideological benign transformation in society, would use the personal computer as a tool for broad intervention and generalized freedom. As Fred Turner (2008) recounts, a countercultural movement emerged in the US, California, during the period from the mid-sixties to the late seventies, just to be later absorbed by the capitalist industrial and consumer logic by the late nineties. The movement acquired momentum owing to the efforts of Stewart Brand and the Whole Earth and what Turner dubbed «an extraordinarily influential group of San Francisco Bay area journalists and entrepreneurs» gathered in 1985 «on what would become perhaps the most influential computer conferencing system of the decade, the Whole Earth 'Lectronic Link, or the WELL. Throughout the late 1980s and early 1990s, Brand and other members of the network (...) became some of the most-quoted spokespeople for a countercultural vision of the Internet. In 1993 all would help create the magazine that, more than any other, depicted the emerging digital world in revolutionary terms: *Wired*». (Turner, 2008, p.3)

This counterculture current presumed that the computer was a tool of universal communication, offering the possibility of democratic participation to every citizen on Earth, thus providing a way for universal harmony among people. However, as Dieter Mersch pointed out in his digital criticism (Merch, 2017), this

countercultural movement would transit from an idealized revolutionary view of reality to a blundered failure. Mainly because there was an underestimation of the industry interests in the digital new environment:

« (...) the idealisms that transitioned from Herbert Marcuse's "great refusal" into the online communities were founded in the fantasy that the PC might advance to a dissidence machine par excellence, to an agent of the unmasking of antidemocratic and authoritarian formations and transform itself into a new and unimagined critical "tool" (...) that would be capable of finally overcoming social structures based on classical hierarchies. But nothing more has emerged from this for years than (...) circulating slogans like "Don't be evil" or "Making the world a better place" ». (Merch, 2017)

With its own weight in the historical development of digital technologies, the advent of the "resold reader's attention" to publicity sponsors in newspapers, happened as early as 1833 in New York. Although it happened before the cyberculture movement, it was to be a critical element for the implementation of algorithmic reason and, most specially, for the surveilling capitalism methods. Advertisement logic has pervaded Algorithmic rationality penetrating deep in each user intimacy, making the user itself or, more exactly, his attention, the transactable good. Fate has it that users attention business model started with an ambitious entrepreneur's insight, named Benjamin Day, as Tim Wu (2016) recalls:

«Day's idea was to try selling a paper for a penny—the going price for many everyday items (...) But what made the prospect risky, potentially even suicidal, was that Day would then be selling his paper at a loss. (...) He would instead rely on a different but historically significant business model: reselling the attention of his audience, or advertising. What Day understood (...) was that while his readers may have thought themselves his customers, they were in fact his product». (Wu, 2016, pp. 11-12)

This goal implied harvesting readers attention by the most efficient means possible e.g., using sensational news. Of course, the same goal would later become the golden rule for social media giants, now using the most efficient algorithmic means possible. At the turn of the last century, we have witnessed a Copernican revolution in rational thought. Quite simply, computers using symbolic or connexionist frameworks, have displaced aspects of our rationality into the wild. One may talk about a dispersed human rationality, in which differentiated very specialized, and amplified rational attributes of human thought have been emulated and turned back at us.

In an optimistic sense, this extended end-result could be interpreted as a technological fulfillment of Teilhard de Chardin noosphere, however, such an environment, with all its techno-procedures, seems to be of a more mechanistic and totalitarian nature than a spiritual one. In fact, a deployed semi-rationality, of a strictly inferential nature, highly optimized for temporal efficiency, merged with the capitalist shareholder's profit logic, producing what Shoshana Zuboff (2019) termed a "surveillance capitalism".

This situation has not come without regret and self-criticism from some of the engineers initially involved, as stated in the "social dilemma movement" (Orlowski, 2020; Lanier, 2018). Third party critics about this highly intrusive algorithmic environment have also surfaced elsewhere (Vaidhyanathan, 2011; Christakis and Fowler, 2011; Forrohar, 2019; Pasquale, 2016).

It must be emphasized that this semi-rationality, acting outside the human sphere, is not to be thus far mistaken for the human rationality pervaded by an experiential wholeness, frequently besides rational awareness towards a localized goal. This algorithmic rationality, expressing itself by means of a computer, is fragmented and a caricature of the human mind. As a human Go player confronts a winning machine, she drags with her the burden of all hurries that make us human: if there will be enough money to pay the bills and so on. The lack of local attention upon such matters do not make them go away, necessarily affecting overall performance compared to the computer. The computerized semi-rationality goes even further being able to perform a cutting incision that separates rational result from its rational author, as one can appreciate in AI music composition or AI painting using generative deep learning neural networks (ART AI and Flow Machines websites). Algorithmic rationality, the one implemented in AI, must not therefore be mistaken for the idealistic rationality that comes with humanist thought. This is of a critical judging nature. The same judgement capability that endured, as humans imagine, and hope for more just and ethical social structures even if such structures in many ways ended up not being perfect. The computer, with a limited rationality, lacks the critical spirit that arises from an emotional rebellion towards whatever inadequate situation is in place. Unlike Socrates or Descartes, a computer – or better still – an automated reasoning system, cannot doubt itself, and hence explain or justify itself to others or to itself. It seems to lack the possibility for an intellectual craving and a longing for better things.

Should it be possible, however, when the singularity arrives, the discussion will be if computers, very much like us, will be entitled to moral rights or responsibilities, as they might have achieved a consciousness with all the inherent properties and qualities (Gunkel, 2012).

Far still from such venturesome state, my aim concerning algorithmic implemented thought is to point out what are the consequences a detached algorithmic rationality can have to our own cultural evolution as a species. I will begin the discussion by defining what I call Algorithmic Reason.

## **2. Algorithmic Reason as an ideology in contemporary society.**

In the last section I have used the expression “algorithmic rationality” to refer to the implementation of human thought in a machine. This mechanistic rationality is to be understood as an uncomplete version of human thought, where each of the human aspects imitated is isolated from all others and applied to the resolution of very strict problems, optimizing for time efficiency. This, I have said, characterizes a fragmented and dispersed rationality outside the human sphere of experience, having been turned at us by capitalist logic.

Before going any further, I wish now to conceptually define what one might identify as the commonsense attitude (including the one held by the industry) towards the use of such algorithmic rationality. The concept I will be using also refers to a critical concern about such use of a fragmented human rationality. In all due right, so far, I have encountered the expression “Algorithmic Reason” in Deiter Merch’s digital criticism (2017) and in Martí Petit paper (2018) where the influence of AI in politics, as well as its necessary regulation, is discussed. I will now present my own definitory view about Algorithmic Reason, concerning its nature and actual implementation.

I propose that Algorithmic Reason is the ideology that computer technology can be applied to any situation, replacing or even generating human intelligence, independently of any critical thought about: i) its logical and scientific soundness (“in Science and Technology we trust”); ii) its ethical value for the final purposes of the algorithmic procedures involved (“Science and Technology will always do us good”) and iii) its teleological value for human endeavor in the long run (“Science and Technology will primarily evolve the human species”), in the context of which one does not care for any other point, besides functionally solving a problem with maximum efficiency (this in fact would be the stance of surveillance capitalist or of any army institution).

Since any situation involving “algorithmic rationality”, as used in the last section, seems to involve the use of “algorithmic reason”, from now on, I wish to use algorithmic reason as a concept that includes algorithmic rationality, that is, the technological implementation of human rationality in a non-human system. Hence, algorithmic reason is to be taken both as an ideological concept and a technological method.

A more succinct way to define algorithmic reason would then be to say that it is the ideology behind computers by which Science and Technology will replace human intelligence; given that *in Science and Technology we must always trust*; that *Science and Technology will always do us good or is always good*, and that *Science and Technology will ultimately evolve the human species*. The former allegations are taken for granted in the urge to solve practical problems.

It should be mentioned that algorithmic reason, both as an ideology and as a practice, has been implemented as a functional automated in practically all activities in our contemporary societies. One can nevertheless point out four main areas, given its magnitude and importance. The first, of course, is Economy, namely when it comes to search engines activity, massive online commodity sellers, Social Networks and High frequency trading algorithms operating in stock market (millisecond taken orders to buy and sell auctions). I will not include trading algorithms in my analysis, since algorithmic financial market is yet another matter (Pasquale, 2016; Peters, 2017).

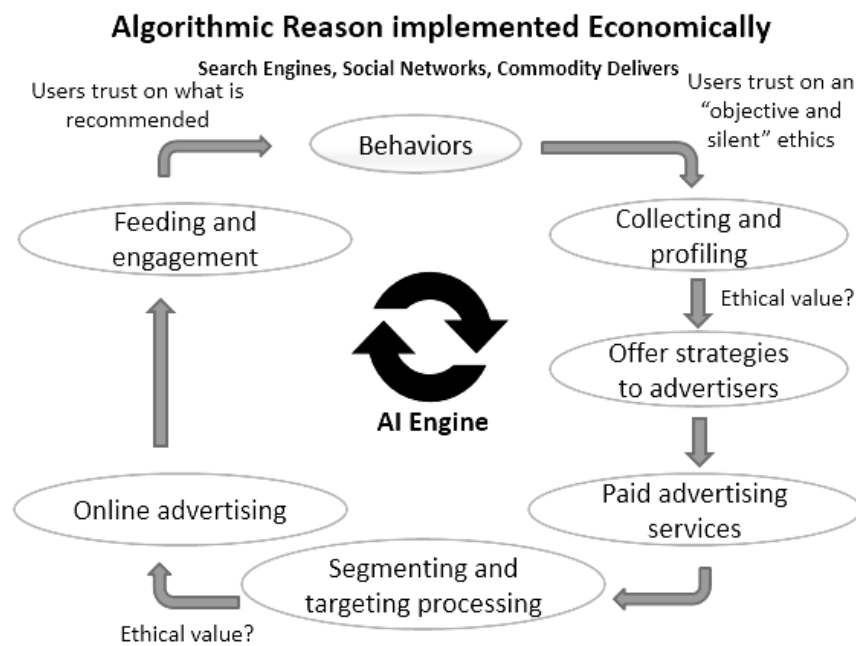
The second main area ceased by algorithmic reason in the twenty first century is human Culture itself, as epistemological conditions for knowledge are being filtered and flattened in the sense that an otherwise enriching global Culture becomes homogenized and globally impoverished, as I will try to show shortly.

Finally, the third critical area colonized by algorithmic reason I wish to refer to, is the political one, namely, when we speak of digital propaganda and political destabilization operated by digital means, mainly through social networks, using fake news and deep fake bot accounts (Woolley and Howard (Eds.), 2018; Benkler and Faris, 2018). As it happens Social Networks pervading the social interaction sphere, will be suggested to be a key component in algorithmic reason implementation, in general, concerning the economic and political dimensions. So let us begin.

Algorithmic Reason implemented Economically starts with an online proposal. Picture, if you will, an “automated wizard” approaching you and saying: “I will offer you everything you ever desired, provided you

let me watch you round-the-clock and record everything you do, think and feel. Will you accept my cookies, as to better serve your preferences / your online experience”?

As it happens, users often accept trusting entirely in the system or just ignoring any reasonable mistrust about it, solely in exchange for some offered functionality, otherwise inhibited by the offering party. It is our attention gathering that the system then seeks to algorithmically collect, as to produce, from our own behaviors and the behaviors of similar minded people, a general profile. There will be different general profiles to be presented before third parties, who will pay for a convenient message transmission (commercial advertising or otherwise) as a social resonance is pursued between the contents of the message and the profile type. To hit the target will be the final and sole objective of the entire operation. However, messages will transform the receivers, whose behaviors will then change, most of the time accordingly to the messages content. This enacts a kind of positive cybernetical loop (Wiener, 1988) involving digital systems and its human users (or perhaps, more correctly, “used”), as the following figure tries to depict:



**Figure 1:** Algorithmic Reason implemented economically.

Now, at some point in time, as the profiling feeds itself from its own steering effect upon human minds, the whole situation will eventually end up in a stable (monolithic) state. People will inform the system of things they like, which the system will kept offering them as likeable. The buzz words associated to this kind of technological application are “feeding” and “engagement”. One may even read on the Internet about “programming for engagement”, as the “right feeding” will bring the right message before the right receiver, while the “right engagement” will be assured by all kind of visual strategies to capture the user’s attention for as long as possible. A dopamine rush kind of strategy will then be used to keep the user coming back online from his otherwise insipid offline existence (e.g., to see the rate of approval from others about something recently published by that same user). Evidently, a sort of digital addiction is being kept, but I think that even more serious issues (to put it mildly) arises from the cybernetical loop long run effects. In this case, flattening the user’s overall preferences and restringing them to a limited predictable universe.

And even if one quite cynically chooses not to care about taste evolution, the next example rises larger concerns, as it affects knowledge evolution in human Cultures. It relates to what I call Algorithmic Reason implemented epistemologically. It certainly must have many more aspects as it is generically associated to the organization and retrieval of information. However, I will refer paradigmatically to the PageRank algorithm created by Sergey Brin and Lawrence Page (1998), the founders of Google, to make my argument. The PageRank algorithm is based on the hyperlinks available in each webpage on the Internet. The overall idea is that a page which is hyperlinked (that is, recommended) by many different low rank pages will be higher ranked, since it got the preference of so many people. On the other hand, a page hyperlinked (again, recommended) by a high rank page will, again, be higher ranked, as it was recommended by a higher mind in the ranking world. This, of course, tries to mimic the very human habit of recommending someone or

something, based either on a quantitative criterion (many recommendations) or else on an authoritative criterion (a higher mind); recommended by an expert on such and such topic. Under the characteristic title "PageRank: bringing order to the Web", Brin and Page, in their seminal paper announcing Google, argue that:

«The citation (link) graph of the web is an important resource that has largely gone unused in existing web search engines. We have created maps containing as many as 518 million of these hyperlinks, a significant sample of the total. These maps allow rapid calculation of a web page's "PageRank", an objective measure of its citation importance that corresponds well with people's subjective idea of importance. Because of this correspondence, PageRank is an excellent way to prioritize the results of web keyword searches». (Brin and Page, 1998, p. 109)

However, the problem is that in the real social world people use recommendations from various sources, even contradictory, and thus both the "many recommendation" and the "expert recommendation" criteria are permanently being evaluated by everyone on a much more critical and personal level. In this way knowledge transference or, in other words, information retrieval performed by each person is a much less static process in society than in the mechanistic aseptic world created by PageRank on the Internet. This would be a lesser problem if there were several competing and different information organizing algorithms on the Web and if people would not so easily (and blindly) accept Google's results for any given search. But as it happens, the average user will always use the same search engine, mesmerized by algorithmic reason ideology, thus believing that what is being shown is the most acute and correct answer to his query. This, I think, produces again an epistemological flattening effect of overall knowledge in the world, which in the long run will inflict a cultural impoverishment on the human mind, on our imagination and creativity. Most especially, in the tech fashion orientated minds of younger people, having less critical resources to look at and for information.

I am therefore suggesting that "page ranking" (and any similar procedure to organize and retrieve information) is a knowledge filtering algorithm that uses average sociability and the algorithmic reason ideological (and demagogical) trustfulness of users in digital systems. In the long run, one may even talk of a lack or failure of teleological value for human endeavor brought about by algorithmic reason.

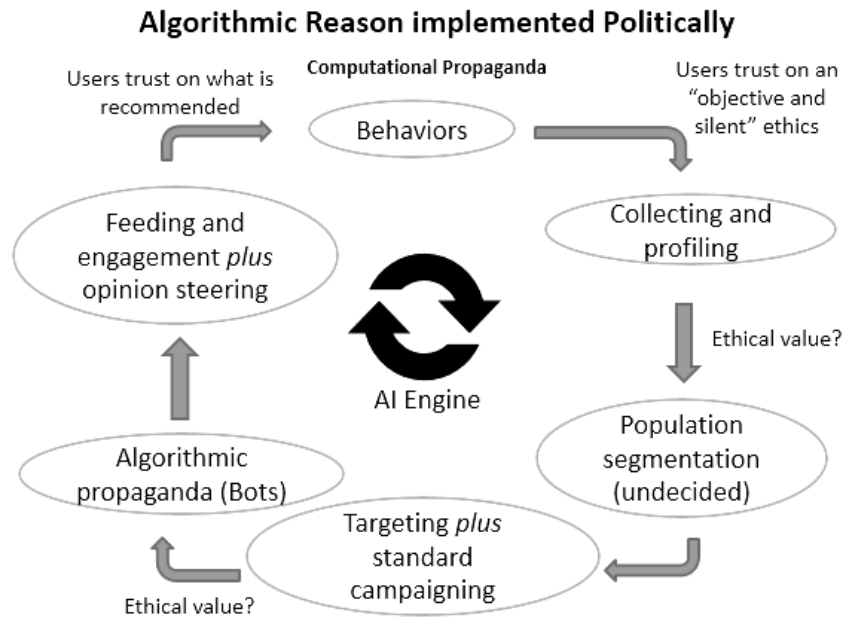
Finally, I will briefly refer to algorithmic reason implemented in the political sphere of human activity. It will be easier, since these days everyone has been informed about online political campaign strategies, from the most benign, to the most vicious, as shown by Jehane Noujaim and Karim Amer in their documentary "The Great Hack" (Amer and Noujaim, 2019) about Cambridge Analytica or Toby Haynes drama film "Brexit: The Uncivil War" (Haynes, 2019).

From a methodological point of view, political implementation of algorithmic reason seems like a particular case of the same structure loop one has when Algorithmic Reason is implemented economically. Of course, the subtil and highly pinpointed targeting that AI engines do at the minds of undecided voters focus, even more, on their emotional response to the right stimulus. These are to be accompanied by the right slogan achieving a change of mind in accordance with the issue or candidate to be voted. The social and political consequences are at this point largely unpredictable, although some of the present effects of what has been recently called "Computational Propaganda" are well known, as reported by Samuel Woolley and Philip Howard:

«Massive social platforms like Facebook and Twitter are struggling to come to grips with the ways their creations can be used for political control. Social media algorithms may be creating echo chambers in which public conversations get polluted and polarized. Surveillance capabilities are outstripping civil protections. Political "bots" (software agents used to generate simple messages and "conversations" on social media) are masquerading as genuine grassroots movements to manipulate public opinion. Online hate speech is gaining currency. Malicious actors and digital marketers run junk news factories that disseminate misinformation to harm opponents or earn click-through advertising revenue.

It is no exaggeration to say that coordinated efforts are even now working to seed chaos in many political systems worldwide. Some militaries and intelligence agencies are making use of social media as conduits to undermine democratic processes and bring down democratic institutions altogether». (Woolley and Howard (Eds.), 2018, p. 3)

Even if this view sounds a bit apocalyptic and could eventually prove to be less serious, I believe that the same mind flattening effect is giving way to a kind of monolithic (totalitarian) state, thus resulting from a similar cybernetic loop as the one applied to economy. The next picture shows the loop schema as I believe it to be implemented politically:

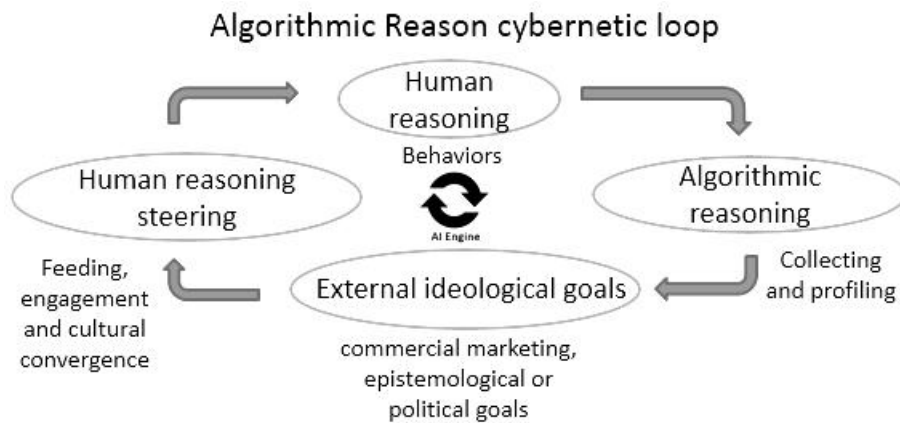


**Figure 2:** Algorithmic Reason implemented politically.

Since it is rather identical to the economical case, I will not extend making it more intelligible than it appears on the figure. Having made my case about the existence of undesired aspects concerning the implementation of algorithmic reason, I will begin instead my analysis about the potential structural faults that I think algorithmic reason carries.

### 3. The three limitations of algorithmic reason.

The examples for algorithmic reason implementation introduced in the last section, can be summarized by what I call the Algorithmic Reason cybernetic loop, portrayed in the overall schema below.



**Figure 3:** Algorithmic Reason cybernetic loop.

Starting from the top middle in the figure, from human reasoning, that is, behaviors, algorithmic reason, comprehended both as a surveilling technology and a (false) reassuring ideology, captures and profiles those behaviors as general behavioral types. Next, vindication coming from external actors (companies' marketers, search engines or political campaigning agents) is injected in the loop, representing all ideological goals to be proposed to the users. Finally, there is content "feeding" and "engagement" trapping (capturing the user's

attention and making her stay online, for as long as possible). All these with the sole purpose of performing culture convergence, defining what I call the “Cultural Steering” effect or the “Human mind Steering” effect.

Before such a scenario, we are now fully entitled to ask if algorithmic reason will (or will not) have a set of undermining properties making it somehow deficient or inherently prejudicial once applied at a worldwide scale. For the purposes of this analysis, let me indeed ask i) if algorithmic reason is logically sound; ii) if algorithmic reason is methodologically sound and iii) if there are motives to doubt its teleological value for human endeavor in the long run.

The first two inquiries are assessing the foundations of the almost “blind” thrust users have in technology, in general, and in AI systems. The presupposed infallibility of digital technology comes, of course, from the high thrust we put, quite righteously let me add, on the scientific method. However, technological application of a scientific achievement does not have to carry with it the same correctness (e.g., in ethical terms), presupposed and indeed practiced along the path to those same results. I will in fact argue for the possible lack of such correctness in contemporary AI by means of the following statements.

My first objection will be of a logical nature. One may consider the Turing Test to be an adequate representation of a suitable type of empirical tests that aim to identify the emergence human intelligence in non-human systems. I have, nevertheless, argued elsewhere (Castro, 2017) that the Turing Test is ultimately undecidable. The reason for this is that the jury performing the test must necessarily prove himself, before another judging party, epistemologically capable of performing the test. That is, it must have been sanctioned to perform the Turing Test (as the judge figure). Of course, this implies that any evoked third-party jury, brought about to confirm the first, must also have been able to justify its authoritative powers before yet another judging party, giving way to an endless logical process. Hence the announced undecidability. Now, if one considers the Turing Test as the best empirical method to identify the presence of an artificial intelligence, if such a method is structurally flawed as may be the case, then there might not be a possibility for a general AI theory. For if there were, one could easily predict the conditions and the experimental ways to test for any AI singularity. This defines my first limitation for algorithmic reason. The second objection I have in mind for AI comes from the following argumentation about its logical soundness.

The performance aspect of AI, that is, its technological implementation as a robust system can also be argued to have the same type of flaw as the one pointed out to the Turing Test. Following the same way of reasoning, one can conclude that a general parsing algorithm – one that checks for the correctness of any other algorithm, in every conceivable situation – does not exist. It would of course be able to check its own general correctness, something that I think would require a kind of structure analogous to an insightful living conscience, for which we do not have (and may not be able to have) a general theory.

This seems to put considerable stress on algorithmic reason as a narrative, presuming, as stated earlier, that computer technology can be applied to any situation, replacing, or even generating human intelligence, independently of any critical thought about its logical and scientific soundness. From an epistemological point of view, algorithmic reason exhibits two major limitations. In one hand, its implementation may not fully replace or generate human intelligence (if not by accident), since a general theory about intelligence may not be a viable epistemological goal. On the other hand, its presumed universal efficiency and correctness cannot be assured since there is no algorithmic way to make an algorithmic system faultless. Both limitations are of a computational origin, leading to epistemological difficulties as asserted.

Finally, the third objection to Algorithmic reason I wish to conjure has to do with its teleological value for human endeavor, at an historical rate. The problem, of course, has to do with the algorithmic reason cybernetic loop main consequence, the Cultural Steering effect and its main consequence, the cultural flattening of values in a homogeneous monolithic state, as already described in the previous sections.

Even if such an effect would not be in place, there would be the obvious ethical problem that algorithms are being used to influence and thus manipulate people, with the rather conspicuous excuse that everyone has agreed to be manipulated. Second, there is the ethical problem that pure algorithmic engineering – as the one applied to capital revenue mechanism in social networks – is not without an ideological and thus ethical responsibility when it comes to its effects in population. The expression “cybernetic loop”, that I have used, alludes both to the theory of Norbert Wiener, as there is an obvious positive feedback mechanism at play, and to his philosophical concerns about the relation between Science, Technology and Ethics. As Steve Heims so eloquently put it:

« (...) Wiener applied his penetrating and innovative mind to identifying and elaborating on a relation of high technology to people which is benign or, in his words, to the human - rather than the inhuman - use of human beings. (...) For him technologies were viewed not so much as applied science, but rather as applied social and moral philosophy». (Wiener, 1988, pp. 11-12)

At the long run, the quite unethical (and totalitarian) use of AI for Human Mind Steering will produce a cultural flattening, smashing cultural diversity and curiosity inquiring, based on a misguided “absolute thrust” in digital contents and the infallible algorithmic way through which they have been produced.

All this does not mean that we should not have AI functioning systems. The main point here is that we must devise the right conditions (and legal regulations) to make artificial intelligence a less invasive tool of human nature, simultaneously embedding it with a kind of digital humanist set of values, which I believe is yet to be formulated.

#### **4. Conclusion: final remarks on algorithmic rationality.**

Artificial Intelligence is becoming a second nature to humankind. This in the sense that it is a corporeal extension of our own rationality. It performs mechanically, owing to a cartesian way of looking at the world, and as such, it is a form of fragmented and highly specialized way of thought that mostly impinges at the resolution of practical problems. I have called this algorithmic rationality and to the ideology favoring it, algorithmic reason.

In the future, humanity may succeed in the creation of a true form of artificial intelligence, that is, an artificial agent ontologically equivalent in all aspects to a human being. If so, such a feat would have been accomplished using the scientific method, which formally can be seen as an instantiation of algorithmic thought, with its axiomatic structures and inferential procedures. Are we to accept that algorithmic rationality can produce an agent capable of a more holistic kind of rational thought, like the one we have in so many other situations in life besides solving practical problems? Given the limitations invoked, this scenario seems rather implausible, if not impossible. Perhaps the result of AI research endeavor will be after all the production of a race of zombie enslaved robots. But if so, are we prepared to plunge culturally in such a social environment? Although we may be optimally served from a functional perspective, will we remain untroubled by the aesthetics of a world comprised of human masters and synthetic slaves?

The main feature of algorithmic thought in non-human agents is the simulation aspect, as the machine makes its best to imitate human nature. It is from the start an empirical deception and accordingly I which to propose that the simulation of an ethical action or scenario (such as the master-slave here portrayed) is still an unethical action or is still an unethical scenario. This obliges us, for instants, to strongly react against violent video games or against the idea that one can have a realistic “amusement” park where it is allowed to “kill” human like figures just because they are artificial (as in the science fiction series *Westworld*). It therefore comes to mind that algorithmic rationality may inherently lack an ethical perspective and that the scientific method may need to be empowered as to include in its methodological inner axiomatic structure ethical procedures, with the same epistemological dignity as the logical ones. Such a revision may perhaps conceal the beginning of a much-needed revolution to deal with the challenges mankind faces in XXI Century.

#### **Acknowledgments**

I would like to thank the Philosophy of AI reading group at CFCUL for the fruitful discussions about AI Ontology and AI Ethics.

I would also like to thank the referee for the useful remarks that helped to improve this article.

This work has been funded by FCT (Portugal) through CFCUL (project FCT UIDB/00678/2020).

#### **References**

- Amer, K., & Noujaim, J. (2019, July 24). *The Great Hack* [Documentary, Biography, History]. The Others.
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press.
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1), 107–117.
- Buhr, M. (2004). *A coragem para a Razão*. In *Razão e Espírito Científico*. Edições Duarte Reis. Lisboa.



- Castro P. (2017). «*Computing Machinery, Intelligence and Undecidability*». *Journal of Theoretical & Computational Science* 4, 160.
- Christakis, N. A., & Fowler, J. H. (2011). *Connected: The Amazing Power of Social Networks and How They Shape Our Lives*. HarperCollins Publishers.
- Coeckelbergh, M. (2020). *AI Ethics*. MIT Press.
- d'Allonnes, M. R. (2018). *La faiblesse du vrai*. SEUIL.
- Dignum, V. (2019). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer.
- Dubber, M. D., Pasquale, F., & Das, S. (Eds.). (2020). *The Oxford Handbook of Ethics of AI*. Oxford University Press.
- Dusek, V. (2006). *Philosophy of Technology: An Introduction*. Wiley-Blackwell.
- Floridi. (2010). *The Cambridge Handbook of Information and Computer Ethics Hardback*. Cambridge University Press.
- Floridi, L. (2013). *The Philosophy of Information*. Oxford University Press.
- Foroohar, R. (2019). *Don't Be Evil: How Big Tech Betrayed Its Founding Principles -- and All of Us*. Currency.
- Gunkel, D. J. (2012). *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. The MIT Press.
- Haynes, T. (2019, January 7). *Brexit: The Uncivil War* [Biography, Drama, History]. Baffin Media, Channel 4 Television Corporation, House Productions.
- Kearns, M., & Roth, A. (2019). *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford University Press.
- Lanier, J. (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. Henry Holt and Co.
- Mersch, Dieter (2017) *Digital Criticism. A Critique of "Algorithmic" Reason*. (n.d.), 2017. Retrieved December 3, 2020, from <http://www.diaphanes.net/titel/digital-criticism-5313>
- O'Neil, C. (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- Orlowski, J. (2020, September 9). *The Social Dilemma* [Documentary, Drama]. Exposure Labs, Argent Pictures, The Space Program.
- Pasquale, F. (2016). *The Black Box Society: The Secret Algorithms That Control Money and Information* (Reprint edition). Harvard University Press.
- Peters, M. A. (2017). *Algorithmic Capitalism in the Epoch of Digital Reason*. *Fast Capitalism*, 14(1), Article 1.
- Petit, M. (2018). *Towards a Critique of Algorithmic Reason. A State-of-the-Art Review of Artificial Intelligence, Its Influence on Politics and Its Regulation* (SSRN Scholarly Paper ID 3279470). Social Science Research Network.
- Turner, F. (2008). *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism* (Illustrated edition). University of Chicago Press.
- Vaidhyanathan, S. (2011). *The Googlization of Everything: (First edition)*. University of California Press.
- Wiener, N. (1988). *The Human Use Of Human Beings: Cybernetics And Society* (New edition). Da Capo Press.
- Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (Reprint edition). Oxford University Press.
- Howard, P. N. (2020). *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. Yale University Press.
- Rushkoff, D. (2021). *Team Human*. W. W. Norton & Company.
- Wu, T. (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads* (First Edition). Knopf.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1st edition). PublicAffairs.
- AI Generated Paintings. (n.d.). ART AI. Retrieved December 3, 2020, from <https://www.artaigallery.com/>
- Flow Machines – AI assisted music production. (n.d.). Flow Machines. Retrieved December 3, 2020, from <https://www.flow-machines.com/>