No Worries – the AI Is Dumb (for Now)

Heimo, Olli I. [1] & Kimppa, Kai K. [1][0000-0002-7622-7629]

[1]Turku School of Economics, University of Turku, Turku, Finland
olli.heimo@utu.fi

Abstract. The hype around Artificial Intelligence (AI) is on. The praise for AI to solve the issues in the society, warnings about AI generating new issues in the society and many ideas on how AI will transform our world upside down are hot topics around both the scientific community and media discussions alike. However, the AI solutions we currently have in use seem to be more or less not worth the hype. Moreover, the current AI solutions in consumer use seem to be lacking the intelligence part. This might not be the case for long, but to understand the possible future scenarios, we must take a look to the current state of affairs. In this paper a current state of consumer AI is discussed and then related to the possible future outcomes of AI development and its effects on society.

Keywords: Artificial Intelligence, AI, Hype, Ethics

1   Introduction

Artificial intelligence is the buzzword for the era and is penetrating our society in levels unimagined before – or so it seems to be (see e.g. Newman, 2018; Branche, 2019; Horaczek, 2019). In IT-ethics discourse there is plenty of discussion about the dangers of AI (see e.g. Gerdes & Øhstrøm 2013, 2015) and the discourse seems to vary from loss of privacy (see e.g. Belloni et al. 2014) to outright nuclear war (See e.g. Arnold & Scheutz 2018; or Hawking and Musk on BBC, Cellan-Jones 2014)) in the spirit of the movie Terminator 2. Yet the current state of AI – what can be seen as a standard user of an information system – seems to be rather different. The AI seems quite dumb.

The AI can be understood in various ways. To make a simplification, in this paper is introduced four different kinds of AI in discourse. First of all, there is the computer game AI, which in fact is a combination of scripts and cheating (i.e. not AI at all) to make the opponents of computer games more lifelike, to make the sensation that you are playing against actual intelligent opponents. This of course is not true because the easiest, cheapest, and thus most profitable way of giving the illusion of a smart enemy is to give the script the power of knowing something they should not. Hence the idea is to give the player the illusion, but the actual implementation is much simpler. That is the art of making a good computer game opponent.

The second one discussed as an AI quite often is data mining, just gathering a lot of information from a huge pile of data. Yet this is usually, and mostly done by scripting; we find patterns, mathematical models and combine tiny bits of data to find similarities, extrordinarities and peculiarities to be analyzed by humans.

Thirdly, we discuss machine learning, mutating algorithms, neural networks and other state of the art AI research. This is something we should actually focus on when discussing AI. These methods make the computer better by every step the computer makes, every decision the computer makes improves the computer, not the user.

The fourth issue often discussed in the field of AI is the living AI, the thinking AI, the "Skynet", the singularity "the moment at which intelligence embedded in silicon surpasses human intelligence" (Burkhardt, 2011). These AIs, sadly, or luckily, depending on the narrative, the utopia and the dystopia, are still mere fiction and in technological scale in the future we cannot yet even comprehend.

To compare anyone of these four with each other is meaningless, because of the fundamental technological levels they rely upon. In this paper the AI discussed is at the third level, the mutating or learning algorithms and the neural networks, which are independent in the sense that they mutate, find connections between things, and are taught, not merely coded to understand the feed and the feedback they are given.

To clarify the term used in this paper, Artificial Intelligence refers to a system which is a mutating algorithm, a neural network, or similar structure where the computer program "learns" from the data and feedback it is given. These technologies are usually opaque (i.e. black box –design), so even their owners or creators cannot know how or why the AI ended up with the particular end-result. (See e.g. Covington, Adams, and Sargin, 2016; Hao, 2019a). As AI has been penetrating the society in many different levels for years, e.g. banking, insurance, and financial sectors (see e.g. Trippi and Turban, 1992, Coeckelbergh, 2015a), the end-product AIs, e.g. Apple Siri, YouTube suggestions, Facebook censorship, Google Ads, etc. seem to be lagging behind. In these cases the AI is pretty predictable – and even stupid.

MIT Technology review (Hao, 2019b) discusses YouTube algorithm with the following topic: "YouTube is experimenting with ways to make its algorithm even more addictive", but yet it can be argued that the addictive nature of YouTube lies in its content, not in its AI, for the AI seems to be suggesting a lot of videos the user has already seen and hardly ever suggests something truly novel. Yet, as in the review is stated, Google is aiming to improve the algorithm to increase the "addictiveness", or as they state "trying to do what it can to minimize the amplification of extreme content". (Hao, 2019b.) For the sake of services provided by the AI, this could be an improvement. For the sake of the users, there might be ethical dilemmas.

Let us also focus on the improvements: currently, if a person starts doing google searches for a product, e.g. a new television, that person surely is in need of a new television and it is obviously beneficial to everybody (even the person themselves) to market different buying options for the said television. Yet when the person stops searching for the televisions there should be two options: either the person does not want a new television anymore or they have already bought one. It should be clear that the marketing potential for the television is diminished. Further still, the google ads keep advertising the television, or any other good the person has done searches for, for

a long time after the searching about the subject has ended and the marketing potential for the person should be directed to other more potential goals. Similar effects can be seen in e.g. Google's YouTube –service where the suggested videos on television reviews will continue long after these videos have searched or Amazon's buying suggestions.

Furthermore, it is a common thing to hear people complaining that there is not much to see in Netflix. Yet, if one compares the start pages of Netflix with each other they can see no overlap at all, because the AI suggests different movies and series to different viewers thus making the catalogue of Netflix seem to be rather small. It is obvious that this is not in the best interests of the company – or the person using the service, as it puts them in a bubble from which it is hard to exit.

2 Dumb AI

Considering how dumb one of the very best corporate AI's– namely the YouTube AI – is there is reason to suspect that we have very little to worry about governmental surveillance AI's either. To clarify, they just cannot be all that much smarter than the YouTube AI, which only seems to be able to handle the last four things the user has been interested in. If the algorithm behind that does not understand to suggest more suitable hits, and it clearly should (if the user subscribes to several dozen channels, surely they want to see more than just the latest finds?), how likely is it that any governmental AI's would be able to consider – in mass surveillance – what we really find important? Similar problems are visible in e.g. Wikipedia's use of bots, which are supposed to clear misinformation, but also seem to delete good faith inputs from anonymous contributors (De Laat 2015a). De Laat also points out that the algorithms used for this purpose are "opaque", i.e. the contributors do not know how they make these decisions to delete – or not (see also Robbins 2018).

The same phenomenon is visible in e.g. Facebook and other Google AI's – they only suggest things that strengthen the bubble in which the users live; and even make a bubble whether the user wants it or not. If the users do not want to stay in the bubble they need to break the bubble by "liking" feeds they do not actually like, but find interesting, and this only aids momentarily until they are re-bubbled in a new similar situation. The same bubble happens in news feeds they follow. To get a populace divided, even though all the information in the world is available in the Internet, we hardly see anything crossing our own views if we only trust the suggestions coming from the sites we follow.

The grandfather of one of the authors already warned about this in the 60s; "follow all the media you can get your hands to, lest you fall prey to your own prejudices." This advice holds even stronger today, with AI's "helping" us to fall into this very trap. He read newspapers from the left, from the right, from in-between – and even tried to find actually independent news. This is getting harder and harder day-by-day. This may be beneficial to the company running the AI – in a short term – but it also alienates the "in-betweens" from the service.

Even with all this evidence, we are worried that the governmental institutions are capable of understanding our actual political preferences – eventually. We must keep in mind that the data, even misprocessed nowadays can be reprocessed when the processors and the AI's get better, even keeping in mind that the current estimates are just that, crude. Although we consider Fleischman's (see e.g. Fleischman 2017) critiques towards general purpose AIs very relevant, as he points out even weak, or narrow purpose AIs which do not reach the level of general artificial intelligence can be powerful – and dangerous – tools for many different purposes.

Also, we need to remember this when we are looking at ads. When we have looked upon a product for a few times and then stop looking at it, for a normal person it could mean two things: we either have bought it or we are not interested in it anymore. However, the AI sees it best to advertise that particular product for months to come.

An excellent example of this is the case of Tay, the learning AI (in this case Learning Software, LS) put forth by Microsoft in 2016. Within a day it was taught to be racist, sexist and anti-Semitic. As Wolf et al. remind us, testing LS like this needs to be very thorough before it is released for use is of utmost importance. (Wolf, Miller & Grodzinsky 2017.)

3   Smart AI

The possibilities for AI penetration are great. Numerous tasks can be improved with AI to be more streamlined, understandable, and efficient. In many cases we can safely assume that some form of AI or data mining algorithm runs tasks in many business sectors, but the businesses understandably want to have some trade secrets on how they run their business. The clear indicator is to follow the money, and e.g. neural networks have been utilised in finance and investment sector for decades (see e.g. Trippi and Turban, 1992). Yet the AI can only be as smart as its creators – if we speak of machine learning and not still sci-fi style AIs. Even now the AI as a machine can do the tasks it is programmed to do and when done correctly with immense efficiency. Thereby to understand why AI is stupid there can be found two main explanations:

1. people are stupid
2. the AI does not have to be smart in this occasion

To focus first on the option of people being stupid: the AI does what is required, but are the people requiring the right things? Do we actually want the AI to continue this streak of non-service or non-profitability, or do we require more out of it? Can we design the requirements better to further improve the business model (or other models) around it?

Secondly, we may accept that the particular AI can be stupid because we do not care. If the engine works, why fix it? It may not be in the best interests of the AI upkeep to put money and effort to the development of a product that still makes the money or the service. This might be the key issue with the aforementioned YouTube AI due to the lack of competition.

Through these points we can derive the possibilities of the future. Whereas the current state remains unclear due to the secrecy among the AI development itself, we can safely assume that these issues will be corrected in the long run. That is, when competition, need for achievement, individual curiosity, or just the basic technological level advances the issues will be solved. YouTube will eventually make a better suggestion bot – or some other service will replace YouTube. The time-frame is the issue. But the AI still requires the commodity of data to develop.

Hence, the data we give now to the AI can be used against us by the next AI, or the one after it. The importance of the data and the classification of it has been understood a while ago and therefore that data can be used to derive different and altogether improved AI systems long after the data has been gathered. This makes the current AI industry dangerous. It is yet unclear at what level the AI can learn from the old data, but it may be safe to assume that it is comparable to new data.

4   Dangerous AI – smart and dumb?

If the AI is smart and in the hands of psychopaths, the AI can be programmed to search our innermost feelings, our deepest desires; that what we want most. Whereas a psychopath, who hones their skill to influence us without mercy, without remorse has human limits, the AI is inhuman. The programming has calculating power gives the AI possibilities to experiment and gather experience simultaneously with hundreds of thousands, millions, billions of people, and billions or trillions of use cases. Moreover when one is using a psychopathic AI, one knows one is using an AI, but one does not know the means of the AI; the AI will learn. Therefore, the psychopathic AI is an AI used by a psychopath, and all corporations (currently) are more or less psychopaths, as they are by law programmed to be so, hence, an AI owned by a corporation is programmed to serve psychopathic needs by a psychopathic programming to get the best result out of everyone. The AI cannot – by default – see you to be more than a means, rather than an end in yourself (Kant 1970/1785). Since this is not a dystopia, this is an as-is situation; this is what the AIs are doing now, we can see what the AI is doing, and if we can see what the AI is doing, it is not that impressive, is it?

But if the AI is dumb, what is the matter with false positives? They still are false positives and interpreted as true results. This may lead to harm for those of us who are victims of the dumb AI: we might not get private sector services (e.g. bank loans) and we might be subjected to public sector services (e.g. social security or intelligence service surveillance). None the less if the AI makes too many bad calls the work process, i.e. those using it, around it (hopefully) takes them into account and the data is reviewed more thoroughly. A smart AI with improved success ratio will be trusted more in the organization and thus is more dangerous for the victims of false positives.

The more automated this kind of work becomes, the less actual humans are deciding these questions – and this will come (see e.g. Loi, 2015) – the wider the consequences of any mistakes made will necessarily be; as Paul R. Ehrlich is claimed to have posited: "To err is human; to really foul things up requires a computer".

Unfortunately, as de Laat (2015b) points out, if humans trust the institutions which use AI, they may also start to trust the AIs the institution uses – even if there is no basis for this trust, as the way the AI functions can be "opaque" (as in the case of Wikipedia, see de Laat, 2015a) or even fully invisible (as is the case of e.g. financial AIs, see e.g. Coeckelbergh, 2015a). As Coeckelbergh (2015b) hints at, we are in danger of becoming the "dumb masters" of the AI which we use, rather than being in control of what happens when the AI is used.

Especially dangerous, both the dumb and the smart AI becomes, when the government starts using them, because the government lacks the accountability of the system (Heimo, 2018, pp. 41–43) and has a wide access to the citizens' information. When the inherit problem of AI with opaqueness is added, the openness and transparency of governance leads to a situation where the fair, just, and equal treatment of citizens is in danger. Therefore, to justify the use of AI in governance, especially in critical governmental information systems (see e.g. Heimo, Koskinen, and Kimppa, 2013), a drastic measures must be taken into accordance before the solutions can be plausible.

Of course we must also be aware of the private sector not to use their AI solutions against the citizens, but this can be more easily covered with proper legislation where the GDPR is a good start.

5   Conclusions

Therefore, it seems that the AI systems of today – those that can be seen to the public – are quite dumb. This of course does not give the whole picture since the effective AI systems can (and in many cases should) be hidden from the public to hide their whole potential. Nevertheless, this does not diminish the possibility the AI gives if done properly. Most of all we must take care of our data protection since the data gathered about us can be used against us later.

AI has potential. Potential for everything and more than the dystopia creators of the current era give it credit for. The timeline is the problem. As we are told that AI can do various things today, what can we see. We can see mutating AIs, neural networks and the faceless corporations behind them trying to get our money, trying to get advertising money from us, trying to make us simultaneously customers, producers, informants and more (for an opposite view, see e.g. EN, I, 1258a8-14). Yet as we see we are doing other things than just giving our time, effort and money to these corporations. They do give us a valuable service and we pay back with our time and money to them, but we still do not seem to be that affected by their AIs. They do not seem to take control of our daily lives with their advertised, even more addictive services. We must give them the credit for the services they provide, but the services given the amount of possibilities and content they deliver could be manifold. What Netflix, YouTube, Twitter, Facebook could offer would be out of the bubble could be out of the bubble they tie one in to the multitude of content they have to offer – and they seriously do have more content than is available in the bubble. The dumb AI keeps the users in the videos the user has already seen, in the music the user has already listened, in the publishers the user has

already read, and in the opinions the user already agrees with. To keep the content interesting the AIs should be tweaked more to find development

In the face of AI hype we often forget that the promises of AI are just that, promises. We may see them as strengths, weaknesses, opportunities or threats, but at the end of this day, as we can see from the current state of the AI we use, they are just promises. If we want to understand what is possible with these kinds of systems we can just follow the money. In addition, the money is in entertainment, the money is in keeping use entertained, keeping us using the system. With the budgets, revenues, incomes at the level of Google, Apple, Facebook, Netflix, Twitter etc. competing for our attention, they should do better. They are good, but they are not that good. We can see through the mistakes their AIs make and thereby they are dumb. A truly smart AI is like a good butler, there, but you cannot see it; serving all the needs you have, and you will miss that good butler only when they are gone. Our attention is someone elses money, and if a system loses our attention, the system's creator starts losing money, or at least stops earning it. Therefore we should not worry, the AI is dumb, and our attention can be directed elsewhere, fortunately, so do not worry, the AI is dumb – for now.

References

1. Aristotle, Nicomachean ethics.
2. Arnold T. & Scheutz M. (2018) The "big red button" is too late: an alternative model for the ethical evaluation of AI systems, Ethics and Information Technology, 20:59—69.
3. Belloni, A. et al. (2014) Towards A Framework To Deal With Ethical Conflicts In Autonomous Agents And Multi-Agent Systems, CEPE 2014.
4. Branche, P. (2019), Artificial Intelligence Beyond The Buzzword From Two Fintech CEOs, Forbes, Aug 21 2019, https://www.forbes.com/sites/philippebranch/2019/08/21/artificial-intelligence-beyond-the-buzzword-from-two-fintech-ceos/#43f741c7113d
5. Cellan-Jones, R. (2014) Stephen Hawking warns artificial intelligence could end mankind, BBC, https://www.bbc.com/news/technology-30290540
6. Coeckelbergh, M. (2015a) The invisible robots of global finance: Making visible machines, people, and places, SIGCAS Computers & Society, Vol. 45, No. 3, pp. 287—289.
7. Coeckelbergh, M. (2015b) The tragedy of the master: automation, vulnerability, and distance, Ethics and Information Technology, 17:219—229.
8. Covington, P., Adams, J., and Sargin, E. (2016) Deep neural networks for youtube recommendations. Proceedings of the 10th ACM conference on recommender systems. ACM, 2016.
9. De Laat, P. B. (2015a) The use of software tools and autonomous bots against vandalism: eroding Wikipedia's moral order? Ethics and Information Technology, 17:175—188.
10. De Laat, P. B. (2015b) Trusting the (Ro)botic Other: By Assumption? SIGCAS Computers & Society, Vol. 45, No. 3, pp. 255—260.
11. Fleischman W. M. (2017) Language Matters: Words Matter, CEPE/Ethicomp 2017.
12. Gerdes, A. & Øhstrøm, P. (2013) Preliminary Reflections on a Moral Turing Test, Ethicomp 2013, 167—174.
13. Gerdes, A. & Øhstrøm, P. (2015) Issues in robot ethics seen through the lens of a moral Turing test, JICES 13/2:98—109.

14. Hao, K. (2019a), We can't trust AI systems built on deep learning alone, MIT Technology Review, September 27, 2019, https://www.technologyreview.com/s/614443/we-cant-trust-ai-systems-built-on-deep-learning-alone/

15. Hao, K. (2019b) YouTube is experimenting with ways to make its algorithm even more addictive, MIT Technology Review, September 27, 2019, https://www.technolo-gyreview.com/s/614432/youtube-algorithm-gets-more-addictive/

16. Heimo, O. I. (2018) Icarus, or the Idea Toward Efficient, Economical, and Ethical Acquirement of Critical Governmental Information Systems, Ph.D. Thesis, Turku School of Economics, University of Turku. https://www.utupub.fi/handle/10024/146362

17. Heimo, O. I., Koskinen, J. S. & Kimppa, K. K. (2013) "Responsibility in Acquiring Critical Governmental Information Systems: Whose Fault is Failure?" Ethicomp 2013.

18. Horaczek, S. (2019), A handy guide to the tech buzzwords from CES 2019, Popular Science Jan 9 2019, https://www.popsci.com/ces-buzzwords/

19. Kant, I. (1970) *Kant on the foundation of morality: a modern version of the grundlegung.* Translated by Brendan E. A. Liddell, Indiana University Press. Other versions also used. Originally *Grundlegung zur Metaphysik der Sitten* , published 1785.

20. Loi, M. (2015) Technological unemployment and human disenchantment, Ethics and Information Technology, 17:201—210.

21. Newman, D. (2018) Top 10 Digital Transformation Trends For 2019, Forbes, Sep 11, 2018, https://www.forbes.com/sites/danielnewman/2018/09/11/top-10-digital-transformation-trends-for-2019/#279e1bca3c30

22. Robbins S. (2018) The Dark Ages of AI, Ethicomp 2018.

23. Trippi, R. R. and Turban, E. (Eds.) (1992). Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance. McGraw-Hill, Inc., New York, NY, USA. https://dl.acm.org/citation.cfm?id=573193

24. Wolf, M. J., Miller, K. & Grodsinsky, F. S. (2017) Why We Should Have Seen That Coming: Comments on Microsoft's Tay "Experiment," and Wider Implications, CEPE/Ethicomp 2017.