AMSTERDAM LAW FORUM
VU UNIVERSITY AMSTERDAM

Commentary

REGULATING ARTIFICIAL INTELLIGENCE: PLEASE APPLY EXISTING REGULATION

Tina van der Linden[*]

I. Introduction

Artificial Intelligence (AI) has been around since the 1950s.[2] Over the past few years a lot has been said and written about the ethical aspects of AI, as awareness rose of the fact that its use does not just bring benefits, like efficiency, but can potentially also cause harm. In the EU, a so-called High-Level Expert Group on AI was put to work, which produced, among other things, the "Ethics Guidelines for Trustworthy AI" in 2019.[3] Following these ethics guidelines the European Commission proposed a Regulation of Artificial Intelligence, published in April 2021.[4]

In this opinion piece I will explain why this proposed AI Regulation is not a good idea. After having established what it is we are talking about, the piece will explain why the use of AI, in my view, should indeed be regulated – but not in the way proposed by the European Commission. The proposed AI Regulation will be briefly introduced, followed by some critical remarks. My proposal is to take existing regulation seriously by applying and enforcing it.

AI is defined in Article 3(1) of the proposed Regulation as:

> "software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."

Annex I lists the techniques through which AI is developed:

> "(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
>
> (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
>
> (c) Statistical approaches, Bayesian estimation, search and optimization methods."

---

[*] *Tina van der Linden is an Assistant Professor in Law, Ethics and Technology at Vrije Universiteit Amsterdam. She has been working in the field of AI and Law for over 30 years. In 1994 she defended her PhD thesis on the theoretical assumptions underlying legal expert systems.*

[2] Rockwell Anyoha, 'Can Machines Think?' (*The History of Artificial Intelligence*, Science in the News, 28 August 2017) <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/> (accessed 27 October 2021).

[3] European Commission, 'Ethics guidelines for trustworthy AI' (Shaping Europe's digital future, 8 April 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (accessed 27 October 2021).

[4] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts' (Communication) COM/2021/206 final.

The definition is particularly broad, and as a result, it proves difficult to think of examples of software that would not fall into its scope.[5] For example, tools that I use in my profession as a teacher/researcher, such as search engines and the electronic learning environment seem to be covered by it, as well as any software used by government agencies to make decisions on citizens (such as those concerning benefits or permits) and by businesses to handle their operations.

II. Why Should the Use of AI Be Regulated?

The famous saying: "technology is neither good nor bad, nor is it neutral,"[6] is very relevant for AI, defined in this broad way. Using AI is indeed not neutral, and influences not only what is done, but also the responsibility, accountability and liability for it, and the power balance between those involved. This can be seen in everyday life. Software that is now being called AI has been in use for some decades, and it is becoming increasingly difficult to compare the interactions that take place now and the way in which they took place without the software that we have become accustomed to. Typically, a human operator is tempted to hide behind the software ('The computer says no – please don't blame me, I cannot help it!'), and for the victim, the person affected, it may be difficult to know why (and even that) the decision was taken in the first place, whom to appeal to, and which arguments to use in doing so.

Moreover, it is well known from literature that the use of AI may affect human rights,[7] such as freedom to share and receive information, equal treatment, privacy, data protection, or access to justice. Thus, while the use of AI may have a very positive, empowering impact, it may also put human rights at risk. We could even say that the use of AI may have an effect on human dignity: the use of AI may undermine our free will (through subliminal manipulation techniques), thereby reducing or eliminating our autonomy, and turning us into mere means to be used for certain ends such as efficiency and profit.[8] Examples include digital nudging techniques that exploit our vulnerabilities and the methods used by social media to make them addictive and keep users glued to the screen for as long as possible – as vividly explained in the Netflix documentary, The Social Dilemma, by Jeff Orlowski.[9] It could even get worse: AI itself may develop into an existential threat to humanity.[10]

Cathy O'Neil convincingly describes some very harmful uses, such as teacher assessment, job applications, and decisions about probation.[11] The example of the algorithm developed by Amazon to review job applicants' resumes is also well known: because it was trained on an

---

5 Raimond Dufour, Josje Koehof, Tina van der Linden and Jan Smits, 'AI or More? A Risk-based Approach to a Technology-based Society' (*Oxford Business Law Blog*, 16 September 2021) <https://www.law.ox.ac.uk/business-law-blog/blog/2021/09/ai-or-more-risk-based-approach-technology-based-society> (accessed 27 October 2021).

6 Kranzberg's First Law. See: Melvin Kranzberg, Technology and History: "Kranzberg's Laws", Technology and Culture, July, 1986, Vol. 27, No. 3, pp. 544-560 <https://www.jstor.org/stable/310538> (accessed 3 November 2021).

7 Council of Europe Commissioner for Human Rights, 'Unboxing artificial intelligence: 10 steps to protect human rights' (May 2019).

8 Marija Slavkovik, Clemens Stachl, Caroline Pitman and Jonathan Askonas, 'Digital Voodoo Dolls' (10 May 2021), arXiv preprint arXiv:2105.02738 (2021), [2105.02738] Digital Voodoo Dolls (arxiv.org)

9 Netflix, 'The Social Dilemma' (2020) <https://www.netflix.com/nl/title/81254224> (accessed 27 October 2020).

10 Nick Bostrom, 'Superintelligence: Paths, Dangers, Strategies' (Oxford University Press, 2014).

11 Cathy O'Neil, 'Weapons of Math Destruction: How big data increases inequality and threatens democracy' (Crown, 2016). See also the following video of her, 'Weapons of Math Destruction' (*Talks at Google*, 2 November 2016) <https://youtu.be/TQHs8SA1qpk> (accessed 27 October 2021).

accurate, existing data set that had more men in high-paid jobs, it preferred men's resumes over women's.[12] It is not hard to see how nasty self-fulfilling prophecies are created, how discrimination of certain groups occurs and how victims are disempowered – if you are not invited for a job interview because your resume was not selected by the algorithm, how could you find out that this occurred at all, and why? Even the people who designed the algorithm may be unable to explain its outcome. It is important to recognise that data on gender, ethnic origin, or religion are not needed for a discriminatory result – groups along those lines may emerge from other, innocent data like shopping habits or places visited.[13] Other examples of harmful uses include credit scoring and other types of social scoring,[14] predictive policing, fraud detection, search engines, and social media, and targeted advertising. Users may be locked up in filter bubbles,[15] and thus remain unchallenged by opposing views. This could potentially lead to social disruption and, eventually, harm to the democratic society.

III. The Proposed AI Regulation

The European Commission's AI Regulation[16] sets rules for trustworthy AI, to make sure that AI in Europe respects our values and rules. A risk-based approach is proposed, distinguishing between unacceptable risk (banned), high risk (subject to obligations), limited risk (just transparency obligations), and minimal risk. In this way, Article 5 of the proposed Regulation prohibits AI that contradicts EU values such as: subliminal manipulation resulting in physical or psychological harm, exploitation of children or mentally disabled people resulting in physical or psychological harm, general purpose social scoring, and remote biometric identification for law enforcement purposes in publicly accessible spaces (with exceptions).

Annex III lists AI systems that qualify as high risk: safety components of regulated products and certain stand-alone AI systems in certain fields.[17] The requirements that high-risk AI systems must comply with are specified in articles 8 to 15. Risk management processes in the light of the intended purpose of the AI system must be established and implemented. They need to make sure that high-quality training, validation, and testing data are used, that documentation and design

---

[12] Jeffrey Dastin, 'Amazon scraps secret AI recruiting tool that showed bias against women' (*Reuters*, 11 October 2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> (accessed 27 October 2021).

[13] The well-known example here is that of the supermarket that found out, based on its data, that a particular woman was pregnant, before she had a chance to tell her parents! See Kashmir Hill, 'How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did' (*Forbes*, 16 February 2012) <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/> (accessed 27 October 2021).

[14] See for example on the social scoring system in China Kshetri, Nir (2020)."China's Social Credit System: Data, Algorithms and Implications" IEEE IT Professional, 22(2), 14-18 <https://doi.org/10.1109/MITP.2019.2935662> (accessed 7 November 2021).

[15] Eli Pariser, 'The filter bubble: How the new personalized web is changing what we read and how we think' (New York: Penguin, 2011).

[16] Strangely, it is called "Artificial Intelligence Act" in its heading – the EU only knows Regulations and Directives, according to Article 288 TFEU.

[17] Namely: biometric identification and categorisation of natural persons; management and operation of critical infrastructure; education and vocational training; employment and workers management; access to self-employment, access to and enjoyment of essential private services and public services and benefits; law enforcement; migration, asylum, and border control management; and administration of justice and democratic processes.

logging features are established, that an appropriate degree of transparency is ensured, and that human oversight, robustness, accuracy, and cybersecurity are ensured.

However, according to the official presentation at the launch of the proposed regulation, most AI systems will not be high risk,[18] and only new transparency obligations for certain AI systems will apply.[19]

IV. Objections

I do welcome the fact that the European legislator acknowledges the risks that the use of AI may pose to society and the desire to prevent them – while at the same time, of course, not stifling innovation.[20] That is always the dilemma that is faced by regulators: on the one hand not all the fruits of technological development are beneficial for society, so society should be protected from harmful effects. At the same time, there is the worry fuelled by the industry lobby, that too much regulation might stifle innovation, chase off investments in technology, leaving the jurisdiction that put the regulation in place in a disadvantaged position compared to others in the technology race.

Many comments and suggestions can be made regarding the proposal. For now, I limit myself to three remarks. Firstly, the approach chosen is a very bureaucratic, procedural one. In order to be allowed to use a "high risk" AI system (which are subject to obligations) a number of procedural boxes need to be ticked. The underlying assumption is that if only certain procedures are followed, we can trust the, by then "compliant", AI systems to be fair when used in practice. From a risk-based perspective, the big risk of this regulation is that very harmful, discriminatory, unfair, disempowering practices involving AI systems could be legitimised, because the AI systems used are theoretically compliant. The compliance of the AI system can then be used as something to hide behind: "compliant AI system says no!".

Secondly, Article 10(3) provides that:

> "Training, validation and testing data sets shall be relevant, representative, free of errors and complete. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons on which the high-risk AI system is intended to be used."

This begs the question, how can we guarantee that a data set be all of these things? According to Article 10(2)(f) possible biases must be taken into account. But then again, how can one ever be aware of biases beforehand? What does bias mean here exactly? If the data accurately reflects the current state of affairs (including skewed distributions of scarce resources along lines of gender, ethnicity, religion), is that biased? And what then would a non-biased dataset look like? It would most likely be inaccurate!

Finally, Article 14 requires human oversight for high-risk AI systems. The well-known "human-in-the-loop" or "meaningful human control" has its own problems. What would incentivise a human to deviate from the AI? Deviation might incur several problems, and thus it might be

---

[18] Lucila Sioli, 'A European Strategy for Artificial Intelligence' (CEPS webinar, 23 April 2021) <https://www.ceps.eu/wp-content/uploads/2021/04/AI-Presentation-CEPS-Webinar-L.-Sioli-23.4.21.pdf> (accessed 27 October 2021).

[19] Article 52 AI Regulation.

[20] European Commission, 'Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence' (21 April 2021) <https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1682> (accessed 27 October 2021).

easier just to follow the system.[21] Then there may be the phenomenon of un-learning: will a human still be able to do the job properly once it has been successfully done by an AI for some time?[22] Furthermore, the requirement of human oversight seems to assume that humans are better at the job – but what if that is not the case? We know, for example, that human judges tend to be stricter in their verdicts before lunch than after.[23]

V. Alternative Approach

Automated systems that are now qualified as "AI" have been used for a long time. And it is not the case that they were used in a legal vacuum. For example, their use by government agencies to decide over citizens (regarding benefits, permits and the like) is governed by administrative law; their use by commercial organisations is governed by contract law (including, possibly, consumer protection law). Moreover, the use of algorithms by social media platforms is limited by freedom of expression and data protection legislation applies to the creation, collection, and processing of personal data by both commercial parties and governments.

The position that I would like to defend here is that it is better to apply (and adapt if that turns out to be necessary) existing legislation to the use of AI in such fields than to create special legislation dedicated to the tool used. We need to take seriously, apply, and enforce administrative law (for example, the requirement that governments be transparent and that decisions affecting citizens be adequately motived), data protection legislation (for example, purpose limitation, restriction on automated decision making, the true meaning of consent), the right to an effective remedy, freedom to receive information, consumer protection, and so on. If we did that, we would not need legislation specifically for AI.

The tools used should be irrelevant; it is what is actually done that should be the object of regulation.[24] And someone should be responsible (and thus liable) for how such tools are used: either because they can influence how the AI works, because they profit from the AI or because they took the decision to use this particular AI in the first place. And, importantly, this person who is responsible and therefore liable should not be able to hide: neither behind a computer program nor behind the fact that the tools used have been approved and are certified to be compliant with the AI Regulation.

VI. Conclusion

While the efforts of the European Commission to mitigate the potential harmful effects of AI are welcomed, the proposed regulation is not the right way to achieve this. Let us not reduce the legal regulation of the use of AI to a formal procedure of ticking boxes and fulfilling administrative requirements. Rather, let us apply and enforce existing law: governments should carefully consider all the interests involved in a decision, government should be transparent, purpose limitation applies to use of personal data by both governments and companies, there shall be

---

[21] Yeheskel Hasenfeld, 'Human Services as Complex Organizations' (Sage Publications, 2009).

[22] Cf the statement by Van den Herik, that after three months of flawless advice, the computer will be the judge, no matter what either of them thinks about this. In: H.J. van den Herik, 'Kunnen computers rechtspreken?' (Inaugural speech, 21 June 1991, Gouda Quint) <https://pure.uvt.nl/ws/portalfiles/portal/1200651/KUNNEN__.PDF> (accessed 27 October 2021).

[23] Hannah Fry, 'Hello World: How to Be Human in the Age of the Machine' (Black Swan, 2018), pp. 57-89.

[24] Bryan Casey and Mark A. Lemley, 'You Might Be a Robot' (2019) 105 Cornell Law Review 287 <https://ssrn.com/abstract=3327602> (accessed 7 November 2021).

equal treatment, privacy, freedom of expression, and freedom to receive information. Indeed, all of this, and more!