THE EUROPEAN UNION'S APPROACH TO ARTIFICIAL INTELLIGENCE REGULATION

Marisa Almeida Araújo[1]

Lusíada University, Portugal
e-mail: almeidaaraujo@por.ulusiada.pt

Abstract

The European Union (EU) is emphasizing regulatory efforts on Artificial Intelligence (AI) and AI-based systems and tools, aiming to not only lead the digital era but also shape Europe's digital (sustainable) future. Either a soft-law approach and hard-law responses to AI are clearly part of the EU's digital strategy. Although the arena in which AI is moving can be treacherous, is also highly seductive and impossible to escape from. Finding a balance between a precautionary approach and promoting evolution and technological development is not always easy, especially when dealing with confronting interests, i.e., different regulatory levels, enforcement and compliance, social and economic asymmetries, and cultural differences, besides the transformation in people's lives and the impact in the future generations.

The paper analyses some challenges – legal and ethical – concerning the use of AI and AI tools. In particular, considering respect for fundamental rights and freedoms, non-discrimination, quality and security, transparency, impartiality, and fairness, and "under user control", and the EU regulatory responses to the risks. The aim of the paper is to offer some insights into the hard and soft law responses of the EU to the topic, and the strategy adopted to AI regulation.

Key words: Artificial Intelligence; Human Rights; Democracy; European Union; Policy; Regulation.

JEL: K380

Introduction

After Powers and Ganascia, we can group AI difficulties into five categories. "[C]*onceptual ambiguities, the estimation risks, implementing machine ethics, epistemic issues of scientific explanation and prediction, and oppositional versus systemic ethics approaches*". (Powers & Ganascia, 2020, p. 29). In each category

[1] Ph.D. Auxiliar Professor, Faculty of Law of Lusíada University, Portugal; Integrated Researcher, Center for Juridical, Economic and Environmental Studies (CEJEA).

risks and implications of AI are tangible, not only to a human rights level, but also to democracy and the rule of law.

In fact, the world has entered a new age in its history and the digital revolution is in its course at full speed. A need to build a strategy for AI is pivotal within a human-centric approach as is pledged by policymakers.

The European Union stands at the forefront of the technological revolution and has been highly aware of the importance of taking the lead and defining the EU's strategy for AI and regulating it. Ursula von der Leyen's political guidelines include a topic on a Europe fit for the digital age, recognizing that "[d]*ata and AI are ingredients for innovation that can help us find solutions to societal challenges* […]". (Leyen, 2019)

A society powered by digital solutions, based on democratic values, has a transformative potential and can emerge as a key driver of innovation, competitiveness, and growth, and solve key challenges.

In fact, the technological revolution led by AI and AI tools is reshaping, namely, industry, the economy, labor, and society in general. Finding a balance between a precautionary principle and encouraging evolution is not always a simple task, especially considering contending interests, disparate regulatory levels, social and economic inequalities, cultural differences, and the effect on the way of life. Not only ours but also the future generation's way of life to whom we are responsible within the intergenerational solidarity.

The growth of *techne* assumes, in the words of Hans Jonas, ethical significance considering the central place it occupies in human purpose like in no other issues. (Jonas, 1984, p. 9) Anticipation has never been such an imperative, especially considering that we are also changing, in fact, the very nature of human actions.

AI's influence is increasing, expanding, and raising legal and ethical concerns. How can society harness the power of AI while ensuring fundamental principles, such as transparency, accountability, and fairness? In an AI era where physical frontiers between States are not barriers, a shared commitment is essential not only to address the risks and ensure the above-mentioned fundamental principles but also to mitigate local asymmetries. The "digital divide" can exacerbate inequalities, bias, and exclusion, and the different "[…] *scale and the power generated by AI technology accentuates the asymmetry between individuals, groups and nations*." (World Commission in the Ethics of Scientific Knowledge and Technology, 2019, p. 8)

A study about an interdisciplinary and broad topic like the one presented in the paper is always a difficult academic quest. This study is fundamentally concerned with the ethical and legal implications of AI and AI-based systems and how human rights in particular, but also democracy and the rule of law, can be affected by it. The paper delves into the legal landscape of AI regulation within the EU, exploring regulatory initiatives that are shaping the approach to the EU's AI governance. The author argues the tension between AI and human rights, and the need

to deliver an ethical AI, identifying core values to establish legal frameworks to evaluate, assess, and regulate AI.

The study used qualitative methods, including legal, ethical, and philosophical analyses, resorting to the work developed by international actors, including ongoing and work-in-progress projects, over the topic and a review of literature related to the subject of this study.

AI challenges and risks – an overview

The digital revolution is taking its course and influencing every aspect of our lives. From our virtual personal assistants and self-driving vehicles to robots (and robotic industry), to health care, economy, labor, health, privacy, information, military, and many other sectors, AI and related technologies are pushing the boundaries fueled by Big Data.

Although recognizing AI can be used to help societies overcome challenges and improve people's lives, it has also a negative side and it can affect human rights, democracy, and the rule of law. A tangible tension between supporting innovation and promoting fairness is palpable and defining a strategy within a human-centric approach, as defined, also, by the EU, is the key aspect.

In the report of Thorbjørn Jagland to the 129th Session of the Committee of Ministers, new challenges to humankind were emphasized: "[...] *how to harness the benefits of the artificial intelligence revolution, while identifying and mitigating its threat to human rights, democracy and the rule of law*». (Jagland, 2019, p. 9)

These risks are not a price to pay, and Michelle Bachelet calls for urgent action to assess the risks. Bachelet concluded that "[…] *until compliance with human rights standards can be guaranteed, governments should implement a moratorium on the sale and transfer of surveillance technology.*" (Bachelet, 2021)

Complex technological advances created and implemented using secretive processes, are harder to evaluate and appraise and less control also means less accountability, both in the results generated and the AI architecture itself. According to Martin Ebers, a model of "*Explainable Artificial Intelligence (XAI)*" is important for everyone who might be impacted by it. (Ebers, 2021)

An example of this tension was addressed in the Wisconsin v. Loomis case. (State of Wisconsin v. Loomis) The Wisconsin Supreme Court analyzed for the first time the use of algorithms and the right to due process. In this case, it was used an algorithm to assess the defendant's recidivism. Loomis's defense argued that, among other issues, they were not able to understand how the algorithm predicted the defendant's recidivism, and, since the AI tool was protected by intellectual property (IP) rights it was not possible to challenge the validity of the result. (Araújo & Meireis, A Century After Plessy v. Ferguson. The Struggle for Racial Equality in the Algorithm Era, 2020)

Although, besides IP rights, the opacity of AI tools can raise transparency issues. A vivid case was presented by ProPublica. The organization analyzed the al-

gorithm "Correctional Offender Management Profiling for Alternative Sanctions" (COMPAS). COMPAS is a recidivism algorithm used in the American judicial system. The report concluded that black defendants were twice as likely as white defendants to be misclassified as a higher risk of violent recidivism. (Larson, Mattu, Kirchner, & Angwin, 2016)

Not assuming or concluding that the algorithm is profiling defendants by race, the truth is that the analysis found that black defendants were more likely than white ones, to be incorrectly judged at higher risk of re-offense. (Araújo & Teixeira, Vis-à-vis Artificial Intelligence, Human Rights, Democracy, and the Rule of Law, 2022)

In this matter, the European Commission for the Efficiency of Justice (CEPEJ) addressed the potential risks of discrimination. CEPEJ concluded that AI tools can, in fact, reproduce inequalities and bias in the criminal justice system, and we need to be aware of it, an assess them. By not doing so, instead of correcting it, we can end up legitimizing them. (European Commission for the Efficiency of Justice, 2018, p. 55)

In fact, algorithms depend on a data diet so, the quality of the data is imperative to erase any bias, stereotypes, or data fundamentalism that will, in that case, be reproduced by AI and enhanced, negatively impacting people's lives.

Following the issue, we can also conclude that not only the quality of data is imperative, but also diversity.

The lack of representativeness can push many into limbo. This topic in particular is addressed by the European Parliament. In the Resolution on AI from 2021, the Parliament underlines that many algorithmically driven identification technologies can "[…] *disproportionately misidentify and misclassify and therefore cause harm to racialized people, individuals belonging to certain ethnic communities, LGBTI people, children and the elderly, as well as women*; […]". (European Parliament, 2021)

For this matter, authors like Sacha Costanza-Chock explain that AI encodes norms, values, and assumptions that, if not diverse, will erase many within the margins. (Costanza-Chock, 2018, p. 4) An example that highlights the issue is automatic gender recognition (AGR). Timnit Gebru argues that seeing gender as a static concept is simplistic and unidimensional (Gebru, 2020, p. 259). This static conceptualization can have negative outcomes for the transgender community for instance.

The UN Human Rights Office of the High Commissioner for Human Rights published information on how effective data gathering might alter people's lives. The UN Assistant Secretary-General for Human Rights, Ilze Brands Kehris specifically stated that the "[…] *lack of data, disaggregated by race or ethnic origin, as well as by gender, age, and other factors, hides the disproportionate impact of certain laws, policies and practices on racial or ethnic groups in all areas of life, […]".* (United Nations Human Rights Office of the High Commissioner, 2022)

The examples demonstrate that the variety of decisions that AI systems can make has a direct influence at the human rights level. The recognition that AI has a substantial impact on people's lives - personally, socially, and politically - motivates not just governments and other international actors, but also technological firms and other stakeholders, including non-profit organizations and academia, to address the need for an ethical framework for AI.

An Ethical AI – some initiatives

There are, already, several efforts to address AI challenges and develop the guidelines and principles for an ethical AI.

For instance, just to name some, the Toronto Declaration led by Amnesty International and digital rights group AcessNow, the Montreal Declaration for a responsible AI was an initiative of the Université de Montréal, the Asimolar AI principles, the Barcelona Declaration, ethical guidelines from the Japanese Society for Artificial Intelligence, or the UNI Global Union Top 10 Principles for Ethical AI.

In November 2021, UNESCO adopted the Recommendation of the Ethics of AI. (UNESCO, 2021) Among its values, it establishes the respect, protection, and promotion of human rights, fundamental freedoms, and human dignity, ensuring diversity and inclusiveness, and living in peaceful, just, and interconnected societies.

At the EU level, in 2018, the European Group on Ethics in Science and New Technologies also published a relevant statement on AI, Robotics, and 'Autonomous' Systems and proposed a set of basic principles and democratic prerequisites, based on fundamental values: human dignity, autonomy, responsibility, justice, equity and solidarity, democracy, rule of law and accountability, security, safety, bodily and mental integrity, data protection and privacy and sustainability. (European Group on Ethics in Science and New Technologies, 2018)

In the White Paper on Artificial Intelligence, it is recognized that AI is a strategic technology, and the European approach to AI is centred on promoting "[…] *innovation capacity in the area of AI while supporting the development and uptake of ethical and trustworthy AI across the EU economy*". (European Commission, 2020, p. 25)

The European Commission addressed the necessary deployment of trustworthy AI as a pre-condition for Europe's future competitiveness and prosperity. (European Commission, 2021)

In 2019, the Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission released its Ethics Guidelines for a Trustworthy AI. In the report, the group establishes three main components for a trustworthy AI.

The Guidelines address the impact of AI on fundamental rights and establish three components for a trustworthy AI: it should be lawful, complying with all applicable laws and regulations, it should be ethical, ensuring adherence to ethical

principles and values; and it should be robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm. (Independent High-Level Expert Group on Artificial Intelligence to the European Commission, 2019)

The document sets key requirements for AI to be deemed trustworthy: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being, and accountability.

Soft and hard law responses to regulate AI at the EU level

AI challenges are being addressed by soft law responses but also recent hard law regulatory measures at the EU level. In fact, the EU is taking the lead in regulating AI and providing a comprehensive legal framework.

The European Parliament adopted several resolutions on AI, e.g. on the matters of ethics (European Parliament, 2020), civil liability (European Parliament, 2020), intellectual property (European Parliament, 2021), criminal law, and AI use by the police and judicial authorities in criminal matters (European Parliament, 2021), and education, culture and audiovisual sector (European Parliament, 2021).

The Parliament also set up a Special Committee on AI in a digital age (AIDA) to analyze the future impact of AI in the digital age on the EU economy, investigate the contribution of AI to business value and economic growth, and analyze the approach of third countries.

Besides this soft law approach, the EU is setting for an intense hard law response to AI.

The E-Commerce Directive (Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') can be considered an example. Or, more recently, the EU General Data Protection Regulation.

In addition, the most recent European Commission Digital Services package addresses many of the issues of the online environment. The Digital Services Act (Directive (EU) 2019/770 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services) and Digital Markets Act (Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC) aims to provide a safer digital space.

However, the most palpable legal effort is the AI Act.

In order to promote investment and innovation in AI, improve governance and enforcement, and support the creation of a single European AI market, the AI Act sets the goal of establishing a legal and unified framework for digital governance.

This framework will be built using a risk-based approach while considering the fundamental rights and the Union's values led by EU lawmakers.

The document's final version is expected by the end of 2023 and is the first regulation on AI. It emerges as a global legal standard and the first comprehensive legal framework for AI.

Within the UE the efforts demonstrate a dedication to a crucial issue, making the EU a global leader in the field, and setting the tone for Member States' national AI initiatives.

The Pittsburgh Statement, which enumerates the fundamental principles of the Union, including respect for human rights, environmental protection, the rule of law, non-discrimination, regulatory transparency, market-based commerce, and the freedom to innovate and have innovations protected, is the basis for the Commission's guiding principle that "*AI should work for people and be the force for good in society*." (European Commission, 2020, p. 25)

Conclusion

The paper underlines the relationship between AI and human rights, and some examples on how AI can replicate and enhance some risks to human rights, alongside with creating new ones.

The necessity to develop a strategy and regulate AI, within an ethical framework, is the way to promote development and technological evolution, while ensuring and respecting fundamental rights and freedoms.

It is essential to identify the guiding principles in order to set ethical standards on AI. The regulatory efforts on AI are imperative, especially in a hard law level.

To effectively preserve human rights, it will also need to promote democracy and the rule of law by giving the required legal means to confront resistance to compliance and accountability, locally and at supra-national level. It is especially crucial in light of the fact that technology and AI systems are a global phenomenon and no national or state frontiers are, by any form, boundaries in the digital arena.

Procedures for oversight, compliance, and responsibility must be established within a legal framework that is founded on ethical standards. It is crucial to defend a trinitarian commitment to three fundamental pillars of democracy, human rights, and the rule of law in the, sometimes difficult, AI arena.

EU's policies and strategy on AI, its soft and hard law responses, and, most particularly, the EU AI Act highlights the need for a uniform legal system that can handle the various policy objectives to uphold peoples' rights and fundamental freedoms.

Funding

This work was financed by national funds by FCT – Foundation for Science and Technology, under the Project UIDB/04053/2020.

References

1. Araújo, M. A., & Meireis, A. (2020). A Century After Plessy v. Ferguson. The Struggle for Racial Equality in the Algorithm Era. *Lusiada Law Review, 23/24*, pp. 195-204.
2. Araújo, M. A., & Teixeira, M. (2022). Vis-à-vis Artificial Intelligence, Human Rights, Democracy, and the Rule of Law. In A. R. Almeida (Ed.), *O direito na sociedade digital.* Lisboa: Universidade Lusíada . doi:https://doi.org/10.34628/ynx7-5w59
3. Bachelet, M. (2021, September 14). Committee on Legal Affairs and Human Rights, Parliamentary Assembly Council of Europe. Retrieved January 3, 2022, from https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=27455&LangID=E
4. Costanza-Chock, S. (2018, July 18). Design Justice, A.I., and Escape from the Matrix of Domination. *Journal of Design and Science*. doi:10.21428/96c8d426
5. Ebers, M. (2021, July 15). Regulating Explainable AI in the European Union. An Overview of the Current Legal Framework(s). In L. Colonna, & S. Greenstein, *Nordic Yearbook of Law and Informatics 2020: Law in the Era of Artificial Intelligence.* Retrieved May 1, 2022, from in: Liane Colonna/Stanley Greenstein (eds.), Nordic Yearbook of Law and Informatics: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3901732
6. European Commission. (2020, February 19). *White Paper on Artifitial Intelligence - a European approach to excellence and trust COM(2020) 65 final.* Retrieved February 15, 2022, from https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
7. European Commission. (2021, April 21). *Communication from the Commission to the European Parliament, the council, the European Ecnomic and Social Committee and the Committee of Regions: Fostering a European approach to Artifitial Intelligence (COM(2021) 205 final).* Retrieved April 2, 2022, from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2021%3A205%3AFIN
8. European Commission for the Efficiency of Justice. (2018). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment.*
9. European Group on Ethics in Science and New Technologies. (2018, March 9). *Statement on Artifitial Intelligence, Robotics and 'Autonomous' Systems.* Retrieved March 22, 2022, from Research and Innovation: https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1
10. European Parliament. (2020, October 20). *Resolution on a civil liability regime for artificial intelligence, (2020/2014(INL)).* Retrieved January 5,

2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html
11. European Parliament. (2020, October 20). *Resolution on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)).* Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html
12. European Parliament. (2020, October 20). *Resolution on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)).* Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_EN.html
13. European Parliament. (2021, October 6). *Resolution on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)).* Retrieved April 15, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html
14. European Parliament. (2021, May 19). *Resolution on artificial intelligence in education, culture and the audiovisual sector (2020/2017(INI)).* Retrieved January 7, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_EN.html
15. Gebru, T. (2020). Race and Gender: Data-Driven Claims about Race and Gender Perpetuated the Negative Biases of the Day. In M. D. dubber, F. Pasquale, & S. Das, *The Oxford Handbook of ethics of AI* (pp. 252-269). Oxford: Oxford University Press.
16. Independent High-Level Expert Group on Artificial Intelligence to the European Commission. (2019, April 8). Ethics Guidelines for a Trustworthy AI. Brussels. Retrieved February 15, 2022, from https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf
17. Jagland, T. (2019, May 16-17). Ready for future challenges – reinforcing the Council of Europe. Helsinki. Retrieved February 14, 2020, from https://rm.coe.int/168093af03
18. Jonas, H. (1984). *The Imperative of Responsability: In Search of an Ethics for the Technological Age.* Chicago: The University of chicago Press.
19. Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016, May 23). How We Analyzed the COMPAS Recidivism Algorithm: . *Propublica*. Retrieved February 16, 2022, from https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm
20. Leyen, U. v. (2019). *Political Guidelines for the next European Commission 2019-2024.* Retrieved March 9, 2022, from https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf
21. Powers, T. M., & Ganascia, J.-G. (2020). The Ethics of the Ethics AI. In M. D. Dubber, F. Pasquale, & S. Das, *The Oxford Handbook of Ethics of AI* (pp. 26-51). Oxford: Oxford University Press.

22. State of Wisconsin v. Loomis. (n.d.). Retrieved March 21, 2022, from https://casetext.com/pdf-email?slug=state-v-loomis-22
23. UNESCO. (2021, November 23). Recommendation on the Ethics of Artifitial Intelligence. Retrieved April 1, 2022, from https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000381137&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_e86c4b5d-5af9-4e15-be60-82f1a09956fd%3F_%3D381137eng.pdf&locale=en&multi=true&ark=/ark:/48223/p
24. United Nations Human Rights Office of the High Commissioner. (2022, February 16). Better data collection bolsters human rights of marginalised people. Retrieved February 16, 2022, from https://www.ohchr.org/EN/NewsEvents/Pages/Data-and-human-rights.aspx
25. World Commission in the Ethics of Scientific Knowledge and Technology. (2019, February 26). Preliminary Study on the Ethics of Artificial Intelligence. Retrieved April 2, 2022, from https://unesdoc.unesco.org/ark:/48223/pf0000367823