Теорія та історія держави і права; історія політичних і правових учень

Theory and History of State and Law; History of Political and Legal Doctrines

UDC 340.1 + 321.7

DOI https://doi.org/10.5281/zenodo.15106615

"What if it contradicts fundamental values?" Once again on artificial intelligence and democracy[1]

Yulia Razmetaeva

PhD in Law, Associate Professor at the Department of Human Rights and Legal Methodology, Yaroslav Mudryi National Law University, Hryhoriia Skovorody Street 77, 61024, Kharkiv, Ukraine;

Researcher, Centre for Multidisciplinary Research on Religion and Society, Department of Theology, Uppsala University, Thunbergsvagen 3B, 752 38, Uppsala, Sweden

E-mail: yulia.razmetaeva@gmail.com

ORCID ID: https://orcid.org/0000-0003-0277-0554

ResearcherID: http://www.researcherid.com/rid/J-7472-2016

Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57215646702

Прийнято: 18.03.2025| Опубліковано: 29.03.2025

Abstract. ***Purpose.*** *This paper examines the impact of artificial intelligence on democratic values, such as freedom, justice and equality, focusing on the challenges*

1This research is part of the project The Jean Monnet Center of Excellence European Fundamental Values in Digital Era, 101085385 – EFVDE – ERASMUS-JMO-2022-HEI-TCH-RSCH. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them.

posed by AI-driven decision-making, algorithmic inequality, and the undermining of trust. It presents a theoretical framework for assessing AI deployment from the perspective of fundamental values, while considering democracy as a value in itself. The study explores whether democratic societies can balance technological advancement with the preservation of fundamental values. ***Methods****. The research employs a theoretical value-oriented approach, analysing foundations that constitute the essence of democratic systems. It includes analysing legal doctrines and existing recent literature to assess the implications of AI in governance, in particular in public decision-making. A number of ethical considerations are integrated to provide a comprehensive evaluation of AI's effects on democracy.* ***Results****. Findings indicate that AI deployment in public processes often poses serious challenges, leading to undermining democratic institutions and weakening the relational trust between citizens and governments. AI-mediated governance reinforce existing social inequalities, while the concentration of AI power among a few entities and a lack of justification risks undermining democratic decision-making. Additionally, competition between democratic and authoritarian states in AI development raises concerns about ethical compromises.* ***Conclusions****. While AI offers significant opportunities, its uncontrolled growth may weaken democratic systems. To prevent this, AI governance must prioritize balanced, value-oriented approach. Combining paradigmatic interdisciplinary research and public discussions is essential to ensure that AI strengthens rather than erodes democratic values.*

Keywords: *artificial intelligence, democracy, fundamental values, trust, public decision-making.*

«А якщо це суперечить фундаментальним цінностям?» Ще раз про штучний інтелект і демократію

Юлія Размєтаєва

К.ю.н., доцентка, доцентка кафедри прав людини та юридичної методології Національного юридичного університету імені Ярослава Мудрого, Григорія Сковороди 77, 61024, Харків, Україна;

Дослідниця, Центр мультидисциплінарних досліджень релігії та суспільства, Теологічний факультет Уппсальського університету, Тунбергсваген 3В, 752 38, Уппсала, Швеція

Електронна адреса: yulia.razmetaeva@gmail.com

ORCID ID: https://orcid.org/0000-0003-0277-0554

ResearcherID: http://www.researcherid.com/rid/J-7472-2016

Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57215646702

Анотація. Мета*. У цій статті розглядається вплив штучного інтелекту на демократичні цінності, такі як свобода, справедливість і рівність, зосереджуючись на викликах, пов'язаних із прийняттям рішень на основі ШІ, алгоритмічній нерівності та підриві довіри. Дослідження представляє теоретичну основу для оцінки розгортання штучного інтелекту з точки зору фундаментальних цінностей, розглядаючи при цьому демократію як цінність саму по собі. Стаття вивчає, чи можуть демократичні суспільства збалансувати технологічний прогрес із збереженням фундаментальних цінностей.* ***Методи****. Дослідження використовує теоретичний ціннісно-орієнтований підхід, аналізуючи основи, що становлять сутність демократичних систем. Це включає аналіз правових доктрин та недавньої літератури для оцінки наслідків штучного інтелекту в управлінні, зокрема в процесі прийняття публічних рішень. Для комплексної оцінки впливу штучного інтелекту на демократію враховано низку етичних міркувань.* ***Результати****.*

Отримані результати свідчать про те, що розгортання штучного інтелекту в публічних процесах часто створює серйозні проблеми, що призводить до підриву демократичних інститутів і послаблення довіри між громадянами та урядом. Управління за допомогою штучного інтелекту посилює наявну соціальну нерівність, тоді як концентрація влади, що стає можливою завдяки штучному інтелекту, серед небагатьох суб'єктів і відсутність обґрунтування ризикує підірвати демократичне прийняття рішень. Крім того, конкуренція між демократичними та авторитарними державами у розробці штучного інтелекту викликає занепокоєння щодо етичних компромісів. ***Висновки****. Хоча штучний інтелект пропонує значні можливості, його неконтрольоване зростання може послабити демократичні системи. Щоб запобігти цьому, управління штучним інтелектом має надавати пріоритет збалансованому, орієнтованому на цінності підходу. Поєднання парадигмальних міждисциплінарних досліджень і громадських обговорень є важливим для того, щоб ця технологія зміцнювала, а не руйнувала демократичні цінності.*

Ключові слова*: штучний інтелект, демократія, фундаментальні цінності, довіра, публічне прийняття рішень.*

Problem setting. Artificial intelligence (AI) technologies are becoming ubiquitous and indispensable, penetrating and being deployed in many processes across both public and private life. While their impact may not always be clearly visible, the changes AIs provokes seem inevitable. These hidden but profound transformations, brought about by artificial agents and systems, affect human agents and systems – including core values and societal structures. This type of impact is only one part of the problem.

Another aspect is the extent to which technology is placed at the centre of the social narrative and largely determines the overall direction of global development. In the pursuit of rationality, efficiency, and logic offered by technologies, fundamental values risk being left by the wayside. On the one hand, this state of affairs is supported by many who promote technocracy, whether due to a sincere belief in its advantages

or the pursuit of economic benefits and political power. In particular, Big Tech companies are concentrating extraordinary power and exerting influence far beyond the private sector while avoiding accountability proportionate to that power. As it was rightly noted, “problematically, these companies can take decisions that directly afford or impinge on the realisation of core liberal-democratic political values, such as privacy, freedom of political association, national security and freedom of speech” [1, p. 148]. On the other hand, the nature of certain technologies, including AI, is such that they may contradict values shared by human beings. As Mark Coeckelbergh outlines in his discussion of how AI undermines democratic values, technology has historically been a component of centralised and non-democratic social orders [2]. Therefore, entrusting technology with shaping the public democratic order may be extremely unwise.

Analysis of recent research and publications. As the serious and broad impact of AI on legal, political, and social systems becomes increasingly apparent, a number of recent studies are looking at approaches that would help mitigate risks while not stifling innovation. In particular, they propose to strongly focus on deploying responsible and trustworthy AI [See, e. g., 3; 4; 5], to act proactively rather than retrospectively [6], to implement the concept of adaptive legal framework [7], to give priority to such approaches as responsible innovation and permissionless innovation [8], and to concentrate on equal technology, in particular to advocate impartial AI while recognise “the indispensable role of women” [9].

A number of research that have appeared in the past few years have addressed the impact of AI on democracy, both the transformative and the devastating ones. Specifically, Pascal D. König and Georg Wenzelburger provide a systemic view of what life with artificial agents might look like in a democratically governed society, focusing in particular on issues of responsiveness and accountability [10]. Eleni Christodoulou and Kalypso Iordanou discuss the negative consequences of using AI in digital media, including the unethical exploitation of human behavior, the increase in disinformation, and ultimately the weakening of democracy [11]. Steven Feldstein emphasises danger of generative AI for democracy as it “poses a new and potentially

harmful wrinkle to mass information consumption" [12, p. 120]. Despite the lively scientific discourse, it is worth revisiting the issue of AI's compliance with democratic requirements from a value-oriented perspective.

Identifying gaps. Describing AI systems as agents without intelligence, Luciano Floridi emphasises that "deploying these systems creates unprecedented challenges in control, coordination, accountability, and governance, requiring innovative approaches to regulation and oversight" [13]. While practical, empirically based approaches may be important in addressing this challenge, it seems worth taking a step back to the theoretical, paradigmatic foundations. What is important to us as humans, and more so as people living together in communities and in society as a whole? What values, principles, and foundations that constitute the essence of our social systems are for us non-negotiable? What if the relentless development and deployment of AI contradicts these foundations? These issues require philosophical and legal analysis that may help to spark scientific debate, engage stakeholders outside academia, and ultimately draw broad public attention to certain implications of AI deployment. To fill the gaps I explore mentioned issues from a value-oriented perspective and propose a coherent theoretical framework that can be applied in further research into both the technological impact on democratic systems and the specific impact of AI. I come up with the hypothesis that the negative impact of AI causes the erosion of fundamental values, destroying key pillars of democracy, and therefore suggest that a cautious approach should be taken to the deployment of such technology, especially in public decision-making and government communication with citizens.

Objectives of the paper. The aim of this paper is to present a theoretical framework for assessing AI deployment from the perspective of fundamental values, with a particular focus on democracy – both as a value in itself and as a system that enables other shared values to flourish. To do so I explore the impact of AI on three democratic values: (1) freedom, (2) justice and (3) equality. In parallel, I explain how their manifest in other values and principles, and why they are particularly important in the digital age. With these examples, I support my main argument that AI development and deployment should remain consistent with fundamental values, even

if it means prioritizing them over stimulating innovation at any cost. I also address the counterargument that competing with non-democratic systems in the pursuit of advanced AI is a compelling reason to prioritize rapid development.

Main findings. The democratic values that AI affects today are, first and foremost, freedom, justice, and equality. On the one hand, they are, so to speak, the basic foundations of democratic systems. On the other hand, they are necessary components, without the proper support of which these systems become fragile. Therefore, it becomes critically important to assess the subtle and often imperceptible undermining of these values that occurs as a result of the design of AI models and the deployment of AI systems.

Freedom is one of the fundamental democratic values, which manifests itself in an array of rights and freedoms, such as individual autonomy, freedom of thought, freedom of expression, the right to participate, etc. All of these freedoms are shrinking under AI-mediated influence. For instance, freedom of expression, that "requires citizens to fearlessly express and share their views" [14, p. 154], is being steadily eroded by algorithmic moderation of speech in the digital environment, selective retrospective censorship that can be carried out with great precision using AI (in effect, rewriting history), and the extensive creation and promotion of artificial content.

It is noteworthy that the impact of AI goes far beyond the digital environment. This is primarily due to its influence on the thoughts and preferences of people who use digital tools. For example, AI facilitates targeted influence on voters before national elections. However, there are many other ways AI can be used to affect freedom. These include offline disinformation campaigns that rely on artificial content, such as deepfakes, precise profiling followed by narrowly targeted influence on opinion leaders (such as public figures), and the creation of a convincing illusion that the range of available choices is much narrower than it actually is.

Disinformation campaigns, for instance, disrupt trust, both between citizens and governments, and between citizens. Such campaigns, according to Jérôme Duberry, "are part of a global power play to reduce the influence of liberal democracies and democratic values in the world" [15, p. 9]. Discussing the unregulated deployment of

AI, he highlights that this technology has “specific characteristics, which imply a higher degree of uncertainty and risk in the citizen–government relation” [15, p. 34]. Remarkably, the deployment of AI is happening faster in the private sphere of life, but its presence in the public sphere is already quite serious, although it is not always taken as seriously as it should be. In particular, the freedom to participate in public processes can only unfold properly when participants are well informed, otherwise the results of their choices will not be the results of informed decisions. At the same time, eroded trust negatively affects how people will perceive even reliable and valid information.

In addition to its destructive impact on trust, AI is well-suited to effectively encourage certain choices and constrain others, which can potentially be done by both private actors (technology owners) and governments. Being embedded in public decision-making by governments, AI affects how reasoned and justified decisions gradually turn into calculated ones, which also affects the value of justice.

Considering another democratic value – the value of justice – it is worth mentioning that it includes among other things, fair public decisions, a fair trial and effective remedies. Criticising AI epistocracy, Sean Donahue argues that it does not gives people to develop records of meaningful moral achievement, while democracy provides adequate opportunity to do so [16]. The way AI is used today not only creates obstacles for humans to develop moral virtues, but it will ultimately prevent AI itself from developing moral virtues when and if it reaches the level of self-awareness, consciousness, and true intelligence. Such a situation is currently hypothetical, but given the pace of technological development, the likelihood of its appearance in the foreseeable future cannot be ruled out.

It should be emphasized that there are many efforts to assess and mitigate the dangers that follow the deployment of AI in public decision-making and democratic processes as a whole. Human oversight and control measures are often proposed, as well as the comparison and evaluation of the outcomes of human and AI decision-making. In addition, as outlined, “management should monitor and review periodically the AI robot’s role in decision-making to further improve situational response” [17]. However, certain risks are rather hidden – for example, the inclination of human

decision-makers to rely on seemingly impartial and accurate AI-produced recommendations.

What is more, public decisions – for example, decisions on new legislation or court decisions of public interest – require proper justification to be considered as aligned with justice. AI-mediated or AI-based decisions lack proper justification; they tend to offer calculations instead of justification, convincing to the general public. As Giles Moss rightly pointed out, “to ensure regulation is well justified, the public must have opportunities to consider, challenge and potentially shape justificatory claims” [18]. One of the problems with AI-based decisions is that they lack a motivational component or replace it with percentages of probability of a certain outcome. Another problem is the general lack of clarity about how AI arrives at certain outcomes, which is especially serious in cases involving deep learning models.

Being part of justice, the right to an effective remedy requires mechanisms that provide adequate protection against violations, redress, and fair compensation if they occur. Discussing obligations on states for state-operated AI systems, including government automated decision-making, Maria O’Sullivan emphasizes that “government automation, particularly in key areas such as social security and law enforcement, poses particular risks given that such systems generally interact with individuals from particular socioeconomic backgrounds” [19, p. 212]. Thus, AI can create a disproportionately discriminatory effect for certain social groups, which contradicts the values of democracy. This brings us to the next issue closely related to justice.

The third democratic value, that may erode due to the impact of AI, is equality. It includes equal opportunities, equality of participation and, with the development of digital technologies, also in some sense equality in the benefits one can derive from innovation. This value also closely connected with the prohibition of discrimination; the latter is undermined by certain AI systems, in particular by risk assessment algorithms.

As Simon Lindgren and Virginia Dignum rightly argued, AI is not universal solution for all social problems, which means that approaches of its deploying should

combine technological, humanities social science perspectives [20]. At the same time, there seems to be a trend to try to achieve equality and non-discrimination through the application of advanced technologies, especially AI. What works well for one area (for example, weather forecasting) may not necessarily work in another (for example, fraud prediction). Therefore, when democratic governments deploy AI in the latter mentioned sphere, vulnerable individuals and groups often fall victim to faulty predictions, which can widen economic, social, and digital divides.

Finally, there is a significant risk of undermining democratic systems through AI used in undemocratic ways. One such way is to produce and reinforce disbelief in democratic values and the importance of democracy as a system. With the help of AI authoritarian governments “push the narrative on social media platforms that liberal democracies are not relevant and resilient, and focus on the challenges they present rather than on the benefits they offer” [15, p. 184]. Such actions fuel debates about whether democracies should disregard their values in the pursuit of the most powerful technology, as it would be disastrous if authoritarian systems, unconstrained by ethical considerations, won the race. Indeed, this presents a strong counterargument to the ‘argument of values’, from the perspective that winning this race is a matter of survival. It is quite likely that authoritarian governments armed with artificial superintelligence (or something close to it) could destroy democracies. However, one should ask whether our democratic systems will remain democratic if we follow the Machiavellian path of justifying the means by the end and discarding values in favor of technological victory.

Conclusions. The rapid advancement and deployment of AI technologies present profound challenges to democratic values, including freedom, justice, and equality. While AI has the potential to enhance various aspects of lives of individuals and functioning of society, its integration into decision-making, public discourse, and governance mechanisms must be approached with caution. The erosion of democratic freedoms, the opacity of AI-driven decisions, and the risk of systemic bias all highlight the need for a value-oriented approach to the embedding of AI into democratic systems.

A crucial question remains: can democratic societies maintain their core principles while competing in a technological race with authoritarian systems? If

democratic states abandon their foundational values in the pursuit of AI predominance, they risk undermining the very principles that define them. Rather than prioritizing unregulated innovation at any cost, AI development must be aligned with the ethical and legal frameworks that safeguard democracy. Ultimately, the challenge is not merely about regulating AI but about reaffirming and reinforcing the values that sustain democratic societies. Therefore, ongoing interdisciplinary research, especially fundamental searches, as well as public debate on them are essential to ensuring that AI serves democracy rather than eroding it.

References

1. Lindman, J., Makinen, J., & Kasanen, E. (2023). Big Tech's power, political corporate social responsibility and regulation. Journal of Information Technology, 38(2), 144–159. https://doi.org/10.1177/02683962221113596.

2. Coeckelbergh, M. (2024). Why AI Undermines Democracy and What To Do About It? Cambridge: Polity Press. 144 p.

3. McGovern, A., Ebert-Uphoff, I., Gagne II, D. J., & Bostrom, A. (2022). Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science. Environmental Data Science, 1, e6. https://doi.org/10.1017/eds.2022.5.

4. Burr, C., & Leslie, D. (2023). Ethical assurance: a practical approach to the responsible design, development, and deployment of data-driven technologies. AI and Ethics, 3(1), 73–98. https://doi.org/10.1007/s43681-022-00178-0.

5. Chander, B., John, C., Warrier, L., & Gopalakrishnan, K. (2025). Toward trustworthy artificial intelligence (TAI) in the context of explainability and robustness. ACM Computing Surveys, 57(6), 1–49. https://doi.org/10.1145/3675392.

6. Allen, D., Hubbard, S., Lim, W., Stanger, A., Wagman, S., Zalesne, K., & Omoakhalen, O. (2025). A roadmap for governing AI: technology governance and power-sharing liberalism. AI and Ethics (Online). https://doi.org/10.1007/s43681-024-00635-y.

7. Lescrauwaet, L., Wagner, H., Yoon, C., & Shukla, S. (2022). Adaptive legal frameworks and economic dynamics in emerging tech-nologies: Navigating the intersection for responsible innovation. Law and Economics, 16(3), 202–220. https://doi.org/10.35335/laweco.v16i3.61.

8. Hemphill, T. A. (2020). The innovation governance dilemma: Alternatives to the precautionary principle. Technology in Society, 63, 101381. https://doi.org/10.1016/j.techsoc.2020.101381.

9. Fraile-Rojas, B., De-Pablos-Heredero, C., & Mendez-Suarez, M. (2025). Female perspectives on algorithmic bias: implications for AI researchers and practitioners. Management Decision. https://doi.org/10.1108/MD-04-2024-0884.

10. König, P. D., & Wenzelburger, G. (2020). Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics. Government Information Quarterly, 37(3), 101489. https://doi.org/10.1016/j.giq.2020.101489.

11. Christodoulou, E., & Iordanou, K. (2021). Democracy under attack: challenges of addressing ethical issues of AI and big data for more democratic digital media and societies. Frontiers in Political Science, 3, 682945. https://doi.org/10.3389/fpos.2021.682945.

12. Feldstein, S. (2023). The Consequences of Generative AI for Democracy, Governance and War. Survival (London), 65(5), 117–142. https://doi.org/10.1080/00396338.2023.2261260.

13. Floridi, L. (2025). AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis. Philosophy & Technology 38, 30. https://doi.org/10.1007/s13347-025-00858-9.

14. Kan, C. H. (2024). Artificial intelligence (ai) in the age of democracy and human rights: normative challenges and regulatory perspectives. International Journal of Eurasian Education and Culture, 9(25), 145–166. http://dx.doi.org/10.35826/ijoecc.1825.

15. Duberry, J. (2022). Artificial intelligence and democracy: risks and promises of AI-mediated citizen-government relations. Cheltenham: Edward Elgar Publishing. 244 p.

16. Donahue, S. (2025). AI rule and a fundamental objection to epistocracy. AI & Society. https://doi.org/10.1007/s00146-024-02175-9.

17. Tóth, Z., c, R., Gruber, T., & Loebbecke, C. (2022). The Dawn of the AI Robots: Towards a New Framework of AI Robot Accountability. Journal of Business Ethics 178, 895–916. https://doi.org/10.1007/s10551-022-05050-z.

18. Moss, G. (2025). Digital Regulation and Questions of Legitimacy. Policy and Internet. https://doi.org/10.1002/poi3.433.

19. O'Sullivan, M. (2024). Artificial intelligence and the right to an effective remedy. In: Artificial Intelligence and International Human Rights Law. Cheltenham: Edward Elgar Publishing. Pp. 196–213.

20. Lindgren, S., & Dignum, V. (2023). Beyond AI solutionism: toward a multi-disciplinary approach to artificial intelligence in society. In: Handbook of Critical Studies of Artificial Intelligence. Cheltenham: Edward Elgar Publishing, 163–172. https://doi.org/10.4337/9781803928562.00019.

Acknowledgments. This research is part of the project The Jean Monnet Center of Excellence European Fundamental Values in Digital Era, 101085385 – EFVDE – ERASMUS-JMO-2022-HEI-TCH-RSCH. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them.