AI and Democracy: Concerns, scenarios and ethical dilemmas

Aikaterini Sampathianaki

To cite this article:

AI and Democracy: Concerns, scenarios and ethical dilemmas

Aikaterini Sampathianaki[1,2]

[1] Translator & PhD Candidate in Technoethics at Panteion University, Greece.
[2] Intern, National Commission for Bioethics & Technoethics, Greece.

✉ katiasabathianaki93@gmail.com

Abstract

If Artificial Intelligence envisages the 4.0 Industrial Revolution and if Technoethics is the multi-disciplinary field that sounds out and discerns the ways our value systems are impacted in the light new technologies, this Article seeks to bring forward opinions voiced on the future of human society, politics and democracy. Is the excessive deployment of AI in both private and public sphere capable of affecting our way of thinking, judging, acting, reacting, making (or delegating) decisions and participating in the res publica? Capitalizing on the field of neuroethics and political science we classify the procedures of human political decision-making, while bringing forward the opinions of techno-optimist and techno-pessimist scholars. Line of arguments ranging from bona fide usage of AI, ethical policy making, enhanced democratic representation down to solutionism and democratic perils of Algorithmic Decision-Making, Echo Chambers, AI biases, and gaps in Accountability, Responsibility, Transparency and Explanation will be presented as a bibliography overview. In the Discussion area paradigms and ethical dilemmas will be outlined for the interest of future research.

Keywords: AI, neuroethics, democracy, political decision-making, governance.

ΤΝ και Δημοκρατία:
Προβληματισμοί, σενάρια και ηθικά διλήμματα

Αικατερίνη Σαμπαθιανάκη[1,2]

[1] Μεταφράστρια & υπ. Διδάκτωρ στο πεδίο Τεχνοηθικής. Πάντειο Πανεπιστήμιο, Ελλάδα.
[2] Ασκούμενη, Εθνική Επιτροπή Βιοηθικής και Τεχνοηθικής, Ελλάδα.

Περίληψη

Έστω ότι η Τεχνητή Νοημοσύνη μετουσιώνει την 4.0 Βιομηχανική Επανάσταση και έστω ότι η Τεχνοηθική αποτελεί τον διαθεματικό εκείνο κλάδο που αφουγκράζεται και διερευνά τον βαθμό στον οποίο τα αξιακά μας συστήματα επηρεάζονται υπό το φως των νέων τεχνολογιών, το παρόν άρθρο φέρνει στο προσκήνιο απόψεις επιστημόνων και ερευνητών αναφορικά με το μέλλον της ανθρώπινης κοινωνίας, την πολιτική και τη δημοκρατία. Είναι ικανή η υπερβολική ανάπτυξη της ΤΝ τόσο στην ιδιωτική όσο και στη δημόσια σφαίρα να επηρεάσει τον τρόπο με τον οποίο σκεφτόμαστε, κρίνουμε, ενεργούμε, αντιδρούμε, λαμβάνουμε (ή αναθέτουμε) αποφάσεις και συμμετέχουμε στα κοινά; Αξιοποιώντας το πεδίο της νευροηθικής και της πολιτικής επιστήμης, ταξινομούμε τις διαδικασίες της λήψης πολιτικών αποφάσεων, ενώ προβάλλουμε τις απόψεις τεχνο-αισιόδοξων και τεχνο-πεσιμιστών μελετητών. Υπό τη δομή βιβλιογραφικής επισκόπησης, παρουσιάζονται επιχειρήματα που κυμαίνονται από την καλόπιστη χρήση της ΤΝ, τον λυσιλογισμό [solutionism], την ενισχυμένη δημοκρατική εκπροσώπηση, έως τους δημοκρατικούς κινδύνους της αλγοριθμικής λήψης αποφάσεων [ADM], τους θαλάμους αντήχησης [echo chambers], τις προκαταλήψεις της ΤΝ και τα κενά στη Λογοδοσία, την Ευθύνη, τη Διαφάνεια και την Εξήγηση. Στο τελευταίο μέρος παρουσιάζονται προτάσεις και ηθικά διλήμματα για μελλοντική έρευνα και δημόσιο διάλογο.

Λέξεις κλειδιά: ΤΝ, νευροηθική, δημοκρατία, λήψη πολιτικών αποφάσεων, διακυβέρνηση.

Introduction and methodology

The 4.0 Industrial Revolution mirrored in Artificial Intelligence [henceforth AI] constitutes an undeniable here-and-now reality, urging modern societies to revisit their standards, value systems and contemplate new governance models to achieve human-machines equilibrium. Are we standing on transformative crossroads where AI takes over democracy giving birth to authoritarian-like regime, or is it safe to say that Democracy and AI are set out on a journey of symbiotic co-existence?

Current concerns of academia are rooted in political philosophy, ethics of technology, governance models, neuroethics and decision-making typology, and the role of AI-induced settings in political discourse and public sphere. New concepts such as Algorithmic Decision-Making, Hybrid Media Systems, Echo Chambers, Bubble Effect and AI biases, Big Data abusive usage, deepfakes and their impact on our citizenship-building procedure are tabled by the techno-pessimist front. Techno-optimist scholars stress the positive role of AI systems in participatory democracy, ethical policymaking, administration and bureaucratic settings.

This is a Technoethics oriented Literature Review intended to discern the latest opinions on *hows* and *ifs* AI algorithms, social media platforms and internet-based systems affect the democratic foundations by grooming public opinion, free will political decision making and civic identity.

Methodologically, we combined narrative and thematic approaches, filtering academic work from political scientists, neuroscientists, behavioral economists, technology institutes and democracy watchdogs to depict both techno-optimist and techno-pessimist views on the future of democracies, while bringing forward various scenarios and recommendations. The Discussion session highlights ethical dilemmas and philosophical questions for future research.

While effort was put to ensure coherence and well-structured pace, this paper inevitably falls short of numerous angles, since this is a dynamically growing field evolving countless experts with fresh research emerging as we write. Given its inherently multidisciplinary nature, technoethics has open-end cognitive and conceptual boundaries, yet to be mapped and delimited.

Neuroethics and Political Decision-Making

Political Decision-Making in Human Societies

If free will of free people is the buttress of democracy, discerning the cognitive basis of political decision-making combined with legitimacy and free elections is primordial. The mechanism of human choice is shaped by individualized contexts, and personal, social and cultural determinations often acting as perception systems, biases and brain heuristics.[1] Thaler and Sunstein reiterate the typology of Kahneman and Tversky (1983) pinpointing the common rules of thumbs governing human judgement and decisions: the heuristics of Anchoring, Availability, and Representation.[2] These modalities function as mental shortcuts and affect our judgements and by extension our political reasoning, especially in democracies where legitimacy is founded on the citi-

---

[1] Braun R. Artificial Intelligence: Socio-Political Challenges of Delegating Human Decision-Making to Machines. Institute for Advanced Studies (IHS), Vienna, 2019, p.13.
[2] Thaler, RH, Sunstein C R. Nudge: Improving Decisions About Health, Wealth, and Happiness, Revised & Expanded edition. Penguin Books, New York, 2009.

zens being the main source of mandate. Here is a cyclic effect: political decisions and outputs are interlinked with citizens and turn back to them in the form of views and preferences.[3] So, effective governance means inputs (e.g. citizens' preferences) been translated into outputs (policies).[4] Yet another factor of democratic discourse is called "hermeneutic element" where citizens should actively and critically interpret information instead of accumulating bulks of data, whereas liberal democracies are often depicted as "a social technology" designated to manage societal complexity.[5]

Neuroethics, Free-Will and Decision-making

Neuroscientist Michael Gazzaniga termed neuroethics as a field that comments on life by means of neuroscience embedded methodology.[6] Issues of perceptions, memory, consciousness, free will and decision-making fall in this scope. Key areas of neuroethics also cover brain privacy and informed consent thus often aligning the field with medical and forensic domains. It also delimits cognitive processes such as memory distortion, particularly the phenomenon of false memories, biases and

perception systems. Our brain tends to reshape memories via a "fit-to-adjust" mechanism to fit the (desired) result. The construction of perception systems is also described by neuroethics as an effort of the human brain to "release capacity" been physically unable to hold on to every information. This property is highly exploitable by the (social) media ecosystem which tends to deploy algorithms to "plant" memories, boost emotional addiction and shape perception systems. Damasio's research reinforces this perception by asserting that emotions are the founding stone of reason and logic.[7]

Free will and the cerebral path to moral choices is yet another contribution of neuroethics; it is argued that moral judgements follow a similar cerebral path to other brain activities: ethical dilemmas are brought forward, filtered and examined and final choices emerge (almost automatically) mainly at the anterior cingulate cortex (ACC) where decisions translate into actions paving the way for "free will" property. Some neuroscientists however, put the notion of "free will" to test. Vilayanur Ramachandran, gives an interesting take on Libet's results[8] arguing that decisions are constantly processed by the nonconscious parts of

[3] Scharpf FW, Governing in Europe: effective and democratic? Oxford University Press, 1999.
[4] Klingemann HD, Hofferbert R, Budge I. Parties, Policies, And Democracy (Theoretical Lenses on Public Policy). Western Press, 1994, p.8.
[5] König PD, Wenzelburger G. Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics. Government Information Quarterly, 2020, 3. https://doi.org/10.1016/j.giq.2020.101489.
[6] Gazzaniga M. The Ethical Brain. Dana Press, Washington, DC, 2005.

[7] Ντινόπουλος Θ. Νευροηθική. Επιστημονικές Εκδόσεις Παρισιάνου, Αθήνα, 2008.
[8] For a detailed report of Libet's experiment see Παπαδόπουλος Β. Νευροηθική: Ηθική και νομική ευθύνη. Το πρόβλημα της ελεύθερης βούλησης υπό το φως των ευρημάτων της νευροεπιστήμης, 2016, p.33-35. https://elocus.lib.uoc.gr/dlib/b/9/5/metadata-dlib-1536919653-758322-19292.tkl.
or: https://www.youtube.com/watch?v=6VZqho-8iJY

the neocortex and solutions/ actions then ascend like *bubbles* to our conscious awareness. Ramachandran introduces the notion of "free won't" - i.e., the power to reject solutions proposed by the nonconscious parts of the neocortex.[9]

The determinism and reductivism theories aside, human behavior results from the interaction of brain functions and is affected by social and cultural conditions. Later paragraphs examine how AI and Algorithimic Decision-Making (ADM) run the risk of neutralizing "social accountability" in political decisions.

AI & Democracy: The Techno-optimist perspective

The social benefits associated with new technological advancements are undeniable when (and if) such apparatuses get ethically designed, based on Research Integrity [RI] and Research Security [RS][10] standardization and aligned with the societal core values. If algorithmic properties are deployed considering public benefit, there are some interesting gains for democracies and citizens: direct cognitive upskilling, innovation, research, investments, new jobs and opportunities, let alone a philosophical and ontological shift. Democracy could use AI to help it become more resilient against authoritarian arrhythmia, blind spots and slippery slopes.

AI Boosting civic representation

By leveraging social media and algorithmic fast-track turnaround of world's news and exchange of opinions, democratic representation and informational autonomy of citizens is improved, thus improving political engagement and healthier decision-making.[11] AI applications lend a hand to disabled persons, remote residents and politically detached citizens, allowing them access to fairer information, transparent political views and more qualitative content engagement.

Paulo Savaget, Tulio Chiarini and Steve Evans argue that AI systems improve civic participation in democracy via open-data and online open-source repositories,[12] while others adds that higher engagement mitigates the citizen's dependency on political representatives' elites.[13][14]

AI enhancing political discourse and citizen's DM

Various scholars argue that if properly trained and ethically designed, AI can boost the "democratic potential" by state-of-art con-

[9] Ramachandran VS, Blakeslee S. Phantoms in the Brain: Probing the Mysteries of the Human Mind. William Morrow and Company, HarperCollins, 1999.

[10] Mollaki V, Ziouvelou X, Giouvanopoulou K, Karkaletsis V. Promoting Research Security through Research Ethics and Integrity practices: recommendations for policy actions, 2025. https://doi.org/10.5281/zenodo.15696984.

[11] Ünver HA. Artificial Intelligence, Authoritarianism and the Future of Political Systems. EDAM, Oxford CTGA & Kadir Has University, 2018.

[12] Savaget P, Chiarini T, Evans S. Empowering political participation through AI. Science and Public Policy, 2019, 46(3):369–380.

[13] Pateman C. Participation and Democratic Theory, Cambridge University Press, 1970.

[14] MacPherson CB. The Life and Times of Liberal Democracy, Oxford University Press, 2012.

tent moderation and mitigation of algorithmic biases.[15] It could also avert hate speech, improve political campaigns, filter deepfakes, social bots and other harmful agents, thus allowing human actors to interact ethically and freely. AI induced social media could uphold the political ethos, strengthen democracy, foster rule of law, fight oppression and discrimination and enhance political mobilization, introducing a new era for human rights movements and other "normative shifts with profound political impacts".[16] The same views are echoed by Sgueo[17] while Battista suggests ethical AI upgrades the efficiency of political decisions.[18]

Policy, regulation and international cooperation

In terms of free and democratic elections, AI induced settings could boost transparency and accountability and truly back up democracies.[19] AI systems and Big Data could yield impressive democratic gains for electorates when policymakers deploy them to ameliorate public administration and e-government, let alone mitigate corruption. Big Data serves democracies when ethically applied in the healthcare, justice or security domains.[20]

Sounding out the alarmist voices, governments, unions and organizations around the globe join forces to prioritize cyber security and AI ethics by establishing Ethics Committees,[21] Councils and by drafting regulations, codes and instruments (soft and hard law) to fortify liberal values, democracies and humanity's set of moral principles from any technological wrongdoing in the future. In parallel,

[15] Wojcieszak M, Thakur A, Ferreira Gonçalves JF, Casas A, Menchen-Trevino E., Boon, M. Can AI Enhance People's Support for Online Moderation and Their Openness to Dissimilar Political Views? Journal of Computer-Mediated Communication, 2021, 26: 223–243. https://academic.oup.com/jcmc/article/26/4/223/6298304

[16] Thiele LP. Politics of Technology-Specialty Grand Challenge. Front. Polit.Sci., 2020, 2.

[17] Sgueo G. BRIEFING (Re-)thinking democracy Digital democracy Is the future of civic engagement online? EPRS | European Parliamentary Research Service, 2020. https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/646161/EPRS_BRI%282020%29646161_EN.pdf. In: Jafarova LA. Political institutions in times of AI, and Ethical Aspects of the Digitalization in Politics. SCIENDO: Polish Political Science Review, 2014, p. 8.

[18] Battista D. Political communication in the age of artificial intelligence: an overview of deepfakes and their implications. Society Register, 2024, 8(2).

[19] Klievink B, Romijn BJ, Cunningham S, de Bruijn H. Big data in the public sector: uncertainties and readiness. Information Systems Frontiers, 2017, 19: 267–283.

[20] Höchtl J, Parycek P, Schöllhammer R. Big data in the policy cycle: policy decision making in the digital era. Journal of Organizational Computing and Electronic Commerce, 2016, 26:147–169. https://doi.org/10.1080/10919392.2015.1125187.

[21] Hellenic Republic National Commission for Bioethics & Technoethics is a pivotal example thereof with its latest Opinions on AI in Education and Preventive Health Analytics https://bioethics.gr/en/opinions%20reports-13/opinion-on-the-artifical-intelligence-applications-in-greek-school-29.04.2025-3222 & https://bioethics.gr/en/opinions%20reports-13/the-applications-of-artificial-intelligence-in-health-in-greece-3175

interdisciplinary approaches emerge to bridge law science and justice -one of the pivotal areas of democratic ecology- with information technology to ensure a safe transition for all stakeholders concerned.

AI & Democracy: The concerns' area

Issues of the Present

Legitimacy, Delegation, Representation

People's legitimacy is the cornerstone of mandate in democratic politics. The ever-growing AI role and the questionable neutrality of "machines" could affect the citizenship-building identity and relations in liberal democracies in three areas: participation, power structures and citizen trust.[22] Some surveys indicate that many citizens around the world entertain the possibility of allowing an AI candidate to run for statehood and even an AI president to undertake the governance[23] by even electing and legitimizing an AI President,[24] meaning that we seek ways to shun corruption, nepotism and bad human judgements. By using the "disappointment" as a key argument, we may be vesting too many powers on

the neutral, clean, clear-cut, fair and firstly appearing on the political scenery algorithms, thus risking the creation of new power centers, also known as "epistemic communities" that could harm cultural and civic identities via a future *commonsense* ground where machines "do it better" and that delegation is permissible at all costs.[25]

Algorithmic Decision- Making [ADM] and solutionism in modern political & statehood settings

An Algorithmic Decision-Making [ADM] system ranges from clearly statistical models and reach applications and techniques of Deep-Learning, a procedure that assigns them more agent-like character. ADM sees political decision-making as one more "cognitive task" that needs to be resolved. This embodies Solutionism the belief that technology (and in our contemporary settings AI) offers turnkey solutions for all our societal, political and bureaucratic problems.[26] Democracy however cannot be reduced to equations and statistical data; politicality, diversity and pluralism seem to resist quantification whereas solutionism risks turning citizens inpatient and willing to delegate more and faster powers to AI and ADM models.[27] Also, in terms of legitimacy, there are three limitations: (1) the lack of a ground truth needed for an optimization process; (2) the

[22] Duberry J. Artificial Intelligence and Democracy: Risks and Promises of AI-mediated citizen-government relations. Edward Elgar, Cheltenham, 2022. In: Fest IC, (book review) Utrecht School of Governance Utrecht University, 2023, p.1.
[23] Carpio A. Is it time to automate politicians? The Economist, Jul 31st, 2018.
[24] Davis D. Is There an AI President in Our Future? That Might Be an Upgrade. Wired, May 18, 2017. https://www.wired.com/2017/05/hear-lets-elect-ai-president/

[25] Antoniades A. Epistemic Communities, Epistemes and the Construction of (World) Politics. Global Society, 2003, 17(1), 21-38.
[26] Morozov E. To Save Everything, Click Here: The Folly of Technological Solutionism. PublicAffairs, New York, 2013.
[27] Jasanoff S, Kim SH. Dreamscapes of Modernity. Chicago University Press, 2015.

fragile link between outcomes to preceding political decisions; and (3) the malleability of decision contexts and public perceptions.[28] Some scholars attempt a comparison between the legitimacy of citizens and their human collective intelligence versus the estimated (or anticipated) AI ultra-intelligence or the Artificial General Intelligence; AI intelligence could erode the human voter's agency reducing citizens to passive recipients of data. Human collective intelligence offers stronger safeguards compared to the narrower ADM. The voters-government relationship and therefore delegation, representation and legitimacy are endangered by the technological determinism: if everything is pre-calculated, pre-processed and simply fed to the electorate, what will voters vote for?[29] Lastly, we should be cautious about the imaginary -a commonsense understanding of the shared vision delegation process- shaped and reproduced by rhetoric and power[30] as such imaginaries often go beyond scrutiny.[31]

Hybrid Media Systems, Echo Chambers and Filter Bubbles

The Hybrid Media System is a term that depicts how the social media platforms mutated from communication, interaction, entertainment, diffusion of cultural products channels to tangible political actors, able to shape political opinion, narratives and impact elections outcome via the control of informational flows and the construction of perception systems.[32] We are looking at politically charged algorithms that affect the future of elections, synthesis of parliaments, public administration settings by the power of the connectivity-culture that allows a channel of carefully designed information (some say computational propaganda) to the benefit or detriment of specific power centers.

Social media platforms and AI algorithms are now seen as "living and breathing political actor"[33] while deploying Machine Learning Algorithms (MLAs) to filter, rank and diffuse information,[34] thus allowing the creation of Filter Bubbles and Echo Chambers both intensifying a closed circuit of information coming the end-user's way, according to their preferences and affiliations. Filter Bubbles and Echo Chambers use the so-called *resonance effect* and the repetition technique. This cognitive

[28] König PD, Wenzelburger G. Between technochauvinism and human-centrism: Can algorithms improve decision-making in democratic politics? European Political Science, 2022, 21:6. https://doi.org/10.1057/s41304-020-00298-3.
[29] Helbing D, Frey BS, Gigerenzer G, Hafen E, Hagner M, Hofstetter Y, van den Hoven J, Zicari RV, Zwitter A. Will democracy survive big data and artificial intelligence? Scientific American 2017, 25. https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/
[30] Braun R. *op.cit.*, p.8.
[31] Harvey D. A Brief History of Neoliberalism. Oxford University Press, 2007, p.24.

[32] Chadwick A. The Hybrid Media System: Politics and Power. Oxford University Press, New York, 2017.
[33] Scholz T. Digital Labor: The Internet as Playground and Factory Routledge, New York, 2012. In: Ünver HA, *op.cit.*, 2018.
[34] Reisach U. The responsibility of social media in times of societal and political manipulation. European Journal of Operational Research, 2020, 291(3):906-917.

fragmentation could weaken political knowledge leading to political alienation and social polarization.[35]

AI Biases and Political Decision-Making

Two interactive experiments held in 2024 sounded out the effects of partisan bias in AI language models on political decision-making.[36] Participants exposed to politically biased models were significantly more likely to adopt opinions and make decisions aligned with the AI bias, regardless of their personal political partisanship. By means of content moderation, under-the-radar data harvesting and profiling techniques biases propagate disparities in content (gender etc.), discriminatory opinions, stereotypes, conspiracy theories and intolerance leaving a door open for societal polarization, racism and political violence.[37] Via "persuasive computing" citizens are nudged to specific political behaviors and judgements, thus raising concerns about the direct involvement of profit-making tech companies in the *res publica*.[38] What is more, political campaigns have undergone extreme

makeover over the last decade thus affecting the voting culture and attitude all over the world.[39]

Deepfakes, Sleeper Social Bots & Political Bots

Media ecology is also bleeding out due to yet another digital apparatus, engineered by specific persons or groups of persons, yearning to disorientate the public opinion or create social uprising – the Deepfakes phenomenon. Deepfakes come with audiovisual tampered content and spread disinformation and conspiracy theories. The Malicious Use of Deepfakes (MUD) is a current social problem putting democratic institutions, international security, diplomacy and future civic societies at real risk.[40]

Sleeper Social Bots are AI agentic entities designed to remain dormant for a designated period prior to becoming active and start spreading disinformation.[41] Such bots apply psychographing and micro-targeting techniques on voters during the pre-election periods that could detrimentally affect free elections and democracy.[42]

[35] Cacciatore, MA, Yeo SK, Scheufele DA, Xenos, MA, Brossard D, Corley EA. Is Facebook Making Us Dumber? Exploring Social media Use as a Predictor of Political Knowledge. Journalism Mass Communication Quarterly, 2018, 95 (2), 404–424.

[36] Fisher J, Feng S, Aron R, Richardson T, Choi Y, Fisher DW, Pan J, Tsvetkov Y, Reinecke K. Biased AI can Influence Political Decision-Making, ArXiv, 2024. https://arxiv.org/html/2410.06415v1

[37] Rozado D. Danger in the Machine: The Perils of Political and Demographic Biases Embedded in AI Systems, Manhattan Institute, 2023.

[38] Helbing et al. *op.cit.*

[39] Tomić Z, Damnjanović T, Tomić I. AI in Political Campaigns. South Eastern European Journal of Communication, 2023, 5.

[40] Pashentsev E. Malicious Use of Deepfakes and Political Stability. Academic Conferences and Publishing International Limited, 2020.

[41] Doshi J, Novacic I, Fletcher C, Borges M, Zhong E, Marino,M C, Gan J., Mager S, Sprague D, Xia M. Sleeper Social Bots: A New Generation of AI Disinformation Bots are Already a Political Threat. University of Southern California, 2024.

[42] Brkan M. Artificial intelligence and democracy: The impact of disinformation, social bots and

Political bots share the same technological and engineering philosophy as sleeper social bots and as we will later see they have been causing some serious political turmoil in Canada and the political decision-making of the citizens, raising concerns on the identification, evidence, attribution and enforcement properties of such algorithmic apparatuses.[43]

Big Data

How can we make a rational and safe link between the Big Data and them, potentially harming democratic procedures? Once designed to enable marketing and consumption techniques Big Data are lately seen in the political scenery: fun and easy-to-use AI applications [trained with gazillions of Big Data] and social media platforms opt for profiling, targeting, shaping political campaigns featuring low transparency and questionable ethics, giving special attention to the critical "undecisive" percentage.[4445] There is quantifiable evidence

that the extensive usage of Big Data in civic procedures creates an alarming drawback for democracies jeopardizing fairness, accuracy and pluralism of views while raising surveillance concerns that are inherently incompatible with democratic values.[46] Moreover, those holding the keys to Big Data centers control political voice and policymaking in various areas of governance while intensifying our concerns for accountability and transparency.

The A.R.T. Problem [Accountability, Responsibility, Transparency]

Can algorithms be truly blamed if they make a mistake, or should we put the blame on the biases uploaded by their coders and developers during the LLM training / alignment procedure? The so-called A.R.T. [Accountability Responsibility Transparency] Problem is interlinked with ADM, and the issues of legitimacy. But why is it so difficult for machines to explain themselves? Do we run the risk of stumbling on the so-called black box? Deep learning procedures deploy probabilistic setups of input nonlinear transformations to generate an acceptable level of output accuracy. If unsupervised, such probabilities end up creating inherent social uncertainties that, by design, make ADM outcomes inscrutable and opaque.

political targeting. Delphi Forum. Interdisciplinary Review of Emerging Technologies, 2019, 2 (2): 66-71.
https://delphi.lexxion.eu/article/delphi/2019/2/4
[43] Dubois E, McKelvey FR. Political Bots: Disrupting Canada's Democracy. CJC Policy Portal, December 20, 2024.
https://cjc.utppublishing.com/doi/pdf/10.22230/cjc.2019v44n2a3511
[44] Costa E, Halpern D. The Behavioural Science of Online Harm and Manipulation, and what to Do about it. The Behavioural Insights Team, 2019.
https://www.bi.team/publications/the-behavioural-science-ofonline-harm-and-manipulation-andwhat-to-do-about-it/
[45] Woolley SC, Howard PN. Automation, Algorithms, and Politics| Political Communication, Computational Propaganda, and Autonomous

Agents. International Journal of Communication, 2016, 10:4882–4890.
https://ijoc.org/index.php/ijoc/article/view/6298/1809.
[46] Mavriki P, Karyda M. Big Data Analytics: Big data analytics in e-government and e-democracy applications: privacy threats, implications and mitigation. Int. J. Electronic Governance,2022, 14:4.

An explanation for any decision made should meet at least one of the following conditions[47]:

- Human-interpretable information (at least not creating new challenges) about the factors used in a decision and their relative weight
- An answer to a counterfactual question.

Lastly, algorithms are usually considered "business secrets" fact which further complicates transparency issues even though certified auditing authorities could resolve this problem, or scrutiny could apply in the blueprint algorithm.[48]

The Future: Towards the rise of new regimes?

Shoshana Zuboff has coined the term "surveillance capitalism" arguing that big tech corporates maximize end-users' content engagement via emotion-triggering content to maximize profits. In this age of surveillance capitalism, digital spaces are used as profit-seeking mechanisms instead of zones of knowledge democratization and civic emancipation.[49] Hacker wonders whether tech companies which run, engineer, deploy and monetize algorithms are willing to find ways to eliminate all the pathogenies or mitigate biases?[50] Could "state surveillance" be simply replaced by "digital surveillance" where human behavior is predictable, and forecasts turn quantifiable?

Howse introduces the term "Algorithmic Feudalism" and Treré the term "Totalitarianism Variants". Capitalizing on the Habermasian model of enclosure and distributionary monopoly, one could say that automation of information systems [including AI], lack transparency and accountability and could mitigate political representation and participation. Drawing on Engels' interpretation of totalitarianism and feudalism, power rests with whoever controls the modes of production, mirroring today's elite of IT leading companies.[51] AI Feudalism involves around the narrative of an AI corporatism system offering protection against chaotic settings. Totalitarian regimes often use technology and science in order to

---

[47] Doshi-Velez et al. *op.cit.* There is a significant debate going on about the Articles 13-15 of GDPR (effective May 25, 2018) and the "right to explanation" concerning the existence, logic and envisaged consequences of automated DM systems combined with the right of the Subject to refrain or decline decisions made by automated systems (Article 22 reference to: Council Regulation 2016/679, arts. 13-15, 22, 2016 O.J.(L119) 1). This debate prompts us to consider that meaningful information methods about how AI systems operate is due if we wish to receive (and therefore exercise our right to) the necessary explanation.
[48] Kavanagh D, McGarraghy S, Séamas K. Ethnography in and around an algorithm. SWG Creativity, Reflexivity and Responsibility in Organizational Ethnography, 2015. https://researchrepository.ucd.ie/handle/10197/7348

[49] Zuboff S. Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. Journal of Information Technology, 2015, 30 (1):75–89. https://doi.org/10.1057/jit.2015.5
[50] Hacker P, Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. Common Market Law Review, 2018, 55(4):1143-1185.
In: Coeckebergh M. The Political Philosophy of AI. Polity Press, 2022.
[51] Ünver HA, *op.cit.*

impose force; technology then is stripped by its "enabler" role and turns into an actor.[52] Another term to depict the same worries is "Machine Totalitarianism"; Ball and Snider argue that in totalitarian settings governors and tech companies develop a symbiotic relationship,[53] whereas Walton & Bhabani comment on the labor precarity, followed by excessive technology dominion.[54]

Responding to Marc Zuckerberg's famous phrase "AI will fix this!",[55] some scholars discern an alarming technochauvinism, namely the belief that societies (and liberal democracies) are flawed and erroneous systems that need constant "debugging" and repair, overlooking the human societal properties of diversity and polyphony.[56]

Country-specific cases

A 2019 survey launched by the Center for the Governance of Change at the Spanish IE University sees more than half of European people been ready to give machines a chance in the next-day governance of their countries.[57] In countries such as Germany and Netherlands more than 30% of the citizens would assign AI the governance. In China the 75% openly favor AI parliamentarians despite the regime's current surveillance and social scoring practices. AI / machine learning in China is embedded in the regime's militaristic narrative and could therefore have serious impact on human rights and civil liberties.[58]

60% of US respondents shun the idea of AI politicians, despite the voters' charted susceptibility to social media propaganda (Cambridge Analytica scandal).

59% of Italy's respondents favor the replacement of humans by AI while in the last elections they were found extremely engaged by TikTok political content.[59] A Dutch survey revealed a two-speed paradox: voters would welcome AI in governance, yet human politicians did not incorporate AI agenda in their

---

[52] Foucault M. Discipline & Punish: The Birth of the Prison, (trans. Alan Sheridan). Vintage Books, New York, 1995.

[53] Ball K, Snider L. The Surveillance-Industrial Complex: A Political Economy of Surveillance. Routledge, New York, 2013.

[54] Walton N, Bhabani S. Rethinking of Marxist perspectives on big data, artificial intelligence (AI) and capitalist economic development. Technological Forecasting and Social Change, 2021, 166(1):120576.

[55] The famous response of Facebook CEO Marc Zuckenberg when asked to give explanations in the 2018 Senate Hearing upon issues of misinformation, hate speech and privacy.

[56] Nemitz P. Constitutional democracy and technology in the age of artificial intelligence. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2018, 376 (2133): 1–14.

[57] Results published in 2021 available at: IE University official webpage results: https://www.ie.edu/university/news-events/news/ie-university-research-reveals-1-2-europeans-want-replace-national-mps-robots/ https://www.cnbc.com/2021/05/27/europeans-want-to-replace-lawmakers-with-ai.html

[58] Cyranoski D. Beijing Launches Pioneering Brain-Science Centre. News, Nature, April 5, 2018. In: Ünver HA, *op.cit.*

[59] Battista D. For better or for worse: Politics marries pop culture (TikTok and the 2022 Italian elections). Society Register, 2023, 7(1).

latest campaigns, showing a low degree of politicization.[60]

Japan has gone one step further: in Tokyo mayoral elections, a candidate called Michihito Matsuda suggested delegating political decision-making, policy implementation and governance entirely to the machines.[61]

Six experiments held in US, Spain and Poland monitor the AI involvement in political decision-making. When it comes to political context, respondents prefer human intervention in most online encounters since humans are seen as more just than AI agents. The study also showcased an 'algorithmic aversion' of public opinion due systemic problems curatorial algorithms feature in terms of construction & deployment.[62]

A qualitative survey showed that in Indonesia's 2024 elections over 95% of Gen Z voters (aged 17-29 years) acknowledge been influenced by AI-induced campaigns via microtargeted and personalized content.[63] In Pakistan, AI curation and deepfake proliferation in

the elections caused filter bubbles, misinformation and led to social and political polarization causing biases and oppression of dissidents.[64]

Canada is yet another interesting case where the political bots created the "astroturfing effect" that caused disorientation and misbalance in the last elections. Political bots initially designed as an administrative tool and a means for journalists to scrap public data, turned into instruments of computational propaganda: their ability of automated accounts creation and interaction with other account users, platforms and datasets allowed them to interfere in the online political discourse causing foggy perceptions to all internet participants.[65]

DISCUSSION

Future scenarios

Scholars' recommendations

When weighing the current bibliography, one cannot come to a safe conclusion on whether AI will harm or assist democracy, the reason why some advocate moderation: AI could be trained and remain as politically neutral as possible to make room for human intel-

---

[60] Morosoli S, Kieslich K, Resendez V, van Drunen M. AI Governance in the Spotlight: An Empirical Analysis of Dutch Political Parties' Strategies for the 2023 Elections, 2024.

[61] Efthymiou IP, Efthymiou -Egleton TW, Sidiropoulos S. Artificial Intelligence (AI) in Politics: Should Political AI be Controlled? International Journal of Innovative Science and Research Technology, 2020, 5.

[62] Wojcieszak et al.m *op.cit.*, p.14.

[63] Febriandy RK, Revolusi P. The Digital Political Revolution: The Impact of Artificial Intelligence (AI)-Based Political Campaigns on Voter Perceptions and Decisions in Generation Z In Indonesia. Jurnal Pendidikan Bahasa, 2024, 11(2):444-458.

[64] Raza A, Waqar AM. Algorithmic Curation in Facebook: An Investigation into the role of AI in Forming Political Polarization and Misinformation in Pakistan. Annals of Human and Social Sciences, 2024, 5, No. 2 (S): 219-232. http://doi.org/10.35484/ahss.2024(5-II-S)22.

[65] Dubois et al., *op.cit.*

ligence to keep making decisions.[66] Another moderate view suggests that since we do not know the future of technology, we should shun the attitude of treating it like a fixed event and trying to remedy for all future events.[67] Skeptics suggest that if an AI-Human symbiotic model is to be fine-tuned in a democracy-oriented manner, we need publicly open procedures for LLM models, because ADM is filtered down to all groups (socialities) affecting relational awareness. Braun suggests politicizing the ADM procedure, namely turning our look not inside the machines, but on the outside where they actually function,[68] a pathway from "polis to technopolis" echoing the work of Hannah Arendt. Civic participation, engagement and inclusion in the development process are encouraged; an "in-progress" mentality must be embraced by all stakeholders while we should also create regulatory sandboxes, responsible research and innovation, research integrity and impact-responsiveness-competence assessment instruments.[69] When it comes to political discourse, agenda setting and pre-campaign information it is argued that public interest should be at the core of ethical

faculty of AI applications and tools used thereof.[70]

Monitoring our "digital well-being" and the impact of technology in our physical, mental and psychological aspects and self-understanding is also recommended[71] combined with education, particularly digital and AI literacy and critical thinking falling in the scope of "user's responsibility"; also the implementation of EU funded projects such as SHERPA, SIENNA and PANELFIT gives hope for the monitoring of human rights agenda, well-being and legislative issues rising from the extensive usage of Big Data. An increase in numbers and power of Ethics Committees and Councils is also highly recommended.[72]

Civic education is also vital in combination with digital literacy to help voters identify and avoid social / political bots and computational propaganda on an early stage. Dubois & McKelvey suggest three policy options for political bots and their astroturfing effect on elections: total ban from social media platforms; establishment of 'bot registries' where stakeholders and owners will have to insert information and comply with standardized require-

---

[66] Makridakis S. The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms. Futures 90, 2017: 46–60. https://doi.org/10.1016/j.futures.2017.03.006
[67] Müller VC. Ethics of Artificial Intelligence and Robotics. The Stanford Encyclopedia of Philosophy, 2020. (Ed. Zalta EN). https://plato.stanford.edu/archives/win2020/entries/ethics-ai/
[68] Ethics guidelines for trustworthy AI. EC. (2018c).
[69] Braun R. *op.cit.*, p.21-23.

[70] Tomić et al., *op.cit.,* p.3.
[71] Burr C, Floridi L. The Ethics of Digital Well-Being: A Multidisciplinary Perspective. In: Burr C, Floridi L (ed) Ethics of Digital Well-Being, A Multidisciplinary Approach. Philosophical Studies Series, 2020: 1–29.
[72] Christodoulou E, Iordanou K. Democracy Under Attack: Challenges of Addressing Ethical Issues of AI and Big Data for More Democratic Digital Media and Societies. Politics of Technology, a section of the journal Frontiers in Political Science, 2021:8.

ments [see DSA, AI Act already enacted in EU area]; stronger Codes of Conduct and stricter Road Maps for social platforms concerning the deployment of political bots and the disclosure obligations thereof.[73]

Some others believe that the "state action doctrine" should be applicable to AI developers and IT stakeholders holding them legally accountable just like public servants are.[74] Another interesting suggestion is to revisit the social contract in a way that fits with the latest AI / algorithmic advancements, introducing the terms of Human-in-the-Loop (HITL) and the Society-in-the-Loop (SITL). This entails drafting an algorithmic social contract (using tools to engineer, develop, program, debug and maintain the systems) where diverse human stakeholders would be mediated by AI models and machines. HITL signifies modeling, simulation and interactive ML (Machine Learning) processes whereas SITL entails the HITL accessing mechanisms to negotiate a value system and monitor the degree of compliance of AI systems with new social agreement and how various stakeholders may be affected.[75]

The issue of explanation…and a solution to the AI accountability gap

Coming back to the challenging area of explanation and accountability of AI systems, scholars propose an apparatus of Legally Operative Explanations: although many consider LLMs to be chaotic in their structure and therefore impossible to provide explanations [black box effect] we should be able to understand a distinction between transparency, namely been aware of the manners and principles a system operates and legally operative explanations, namely straightforward answering to questions posed. This is feasible if we enact two modalities: local explanation and counterfactual faithfulness.[76] On a different note, Lessig's fourth modality on system architecture as a means of regulatory constrain ("constraint of the world as I find it") means that coders' choices in design could prove more impactful in terms of transparency than strict (and often strangulating) regulation.[77]

Other scholars go by the optimization of "Sociodiversity" which is as valuable as biodiversity, fueling resilience of society and democracy to unexpected shocks leaving space for the so-called Cultural Genome Project.[78]

Questions and Techno-ethical Dilemmas

Having traced some of the latest academic voices and trends about the coupling of AI with democratic regimes and the risks for tech-totalitarianism, several questions remain to be handled by governments, politicians and policymakers: Beyond a much-discussed global job losses scenario what other changes is AI likely to cause for public bureaucracies? What

---

[73] Dubois et al., *op.cit.*

[74] Crawford K, Schultz J. AI systems as state actors. Columbia Law Review, 2019, 119.

[75] Rahwan I. Society-in-the-loop: programming the algorithmic social contract. Springer Nature Link. Ethics Inf Technol, 2018, 20:5–14. https://doi.org/10.1007/s10676-017-9430-8.

[76] Doshi-Velez et al. *op.cit.*, p. 13-14.

[77] Bietti E. Assessing principles for the regulation of online content: Lessig's modalities of regulation. Media Laws: Law and Policy of the Media on A Comparative Perspective, 2017. https://www.medialaws.eu/wp-content/uploads/2017/01/1.2017-Bietti.pdf

[78] Helbing et al. *op.cit.*

are the challenges and bottlenecks that civil society encounters when deploying AI systems for political participation? Should we boost decentralized information systems and improve inter-operability and collaborative opportunities via digital literacy? Is the Actor-Network Theory a fit-for-all solution to our existential puzzle?

To the best of our understanding, it is advisable to map which types of AI-induced political participation are to be embraced or avoided. Furthermore, we could turn to smart regulation, digitally literate (and therefore bulletproof) constitutions, equitable resources distribution to avoid digital colonialism, while embedding ethics-by-design into AI architecture and growing long-haul strategic foresight models (including but not limited to national blueprint AI strategies). Finally, we should discern, delimit and shield the domains of national security, secrecy and diplomacy against algorithmic glitches and arrhythmia.

As homo sapiens *organic* societies and *silicon* algorithmic blueprints tend to ontologically converge, a meta-human discourse is unveiled. The scientific community is called upon to draft a roadmap for future generations and democracy: humanities must be revamped and further integrated into technological discourse. Creativity, empathy, reciprocity, diversity, pluralism, trust, solidarity and cooperation should be our guiding light.[79] From a philosophical point of view, if AI offers the gift of virtuality allowing us to contemplate alternative realities, does this give us new po-

litical and civic paradigms? New types of Governance? A Democracy of Things perhaps? Can we train our societies to avoid ethics panics, promote prudent regulation while also leaving space for innovation and research integrity and support informational self-determination?

Computational complexity and ontological myths aside, Artificial Intelligence remains a human creation that needs to be embraced and trained with humanitarian values.[80] Technoethics is a fast-growing research arena that points to the obvious: human-machines symbiosis only makes sense if we revisit the human condition and delve deeper into the meaning of life and human societies. In terms of democratic vigilance and societal awareness, perhaps it is worth spending time and resources *now* to avoid future generations been diagnosed with "civic anosoagnosia" where citizens will be unaware of their democracies been incapacitated.

---

[79] Tsekeris C. Industry 4.0 and the digitalisation of society: Curse or cure? Homo Virtualis, 2018, 1(1): 4–12. https://doi.org/10.12681/homvir.18622

[80] Γιαννακόπουλος Γ. Τεχνητή Νοημοσύνη: Μια Διακριτική Απομυθοποίηση. Εκδόσεις Ροπή, 2021.