

Transatlantic Approach to AI

Policy Brief

*By Dominika Hajdu, Policy Director, Miroslava Sawiris, Senior Research Fellow
Democracy & Resilience Center, GLOBSEC*

Summary

Addressing the use of AI from the perspective of human rights compliance and the protection of democratic principles across the transatlantic space, this policy brief argues for AI systems to be rules-oriented and categorized based on the risk they pose to users and society (drawing on best practices from the EU AI Act). Risk definitions should integrate both human rights and democracy protection elements. It recommends any AI systems that could potentially impinge on human rights and/or be exploited to undermine democracy be deemed high risk, necessitating stringent oversight and regulatory measures. Content recommendation algorithms are identified as a case in point underlining the need for action.

Problem

Discussion around human rights in the digital sphere is a prerequisite to the adoption of regulative measures. But this engagement, all too often, tends to be reduced to a narrow focus on freedom of speech and privacy rights. While consideration of these rights is undoubtedly essential, a reductive approach can lead to policy solutions which may be a) disproportionate in their (in)ability to guarantee other (no less important) human rights including non-discrimination (Article 2 in the Universal Declaration of Human Rights)¹ and b) inefficient in addressing democratic backsliding fueled in online spaces (or sometimes even unwittingly accelerating these processes)². The adoption of such policies, furthermore, can detrimentally affect marginalized communities. The use of biased data-trained AI systems is an illustrative example, with inadequate and inconsistent judicial oversight online perpetuating and further amplifying existing inequalities.³

Purpose

The purpose here is to more robustly re-center discussions pertaining to democracy and human rights in the digital age. A particular focus must be assigned to AI systems ensuring the more comprehensive protection of individual rights and freedoms online and to revitalizing democratic processes and protecting these mechanisms from malign actor sabotage. This recentering, facilitated by the work of the Trade and Technology Council in seeking out common solutions, represents a unique opportunity to broaden discussion beyond the EU and US frameworks standing strictly on their own towards a more shared transatlantic democratic digital space⁴. But any move in this direction will require bold steps. Success, meanwhile, could motivate other

¹ <https://www.un.org/en/about-us/universal-declaration-of-human-rights>

² [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU\(2021\)653635_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf)

³ https://www.brennancenter.org/sites/default/files/2021-08/Double_Standards_Content_Moderation.pdf

⁴ <https://www.globsec.org/publications/10-transatlantic-principles-for-a-healthy-online-information-space-endorse-them-here/>

democratic states to search for synergies in internet governance, an important aspiration in an era of rival non-democratic internet governance models (e.g. China).

Main argument

AI systems and algorithms represent a key piece of the puzzle concerning the development of a human rights- and democracy-centered approach to internet governance. These systems indeed play an increasingly instrumental role in our daily lives including the personalization of social media news feeds and search engine results, digital voice assistants and smart devices and online banking security.⁵

The importance of AI systems and their ubiquity are set to grow exponentially over the coming years. Absent proper safeguards, however, their deployment is already afflicting problems on an unprecedented scale. While AI-powered large-scale surveillance technologies are becoming cheap and readily available, AI ethics is lagging and adequate industry benchmarks and consensus lacking.⁶

The EU's AI Act proposal is an example of an initiative aimed at categorizing and regulating AI systems based on the level of risk they pose to society. The tiered approach employed is a prudent path for regulators on both sides of the Atlantic but a strong emphasis must be placed on the inclusion of human rights and the protection of democratic principles when conducting risk assessment. The unethical use of AI systems deployed in both the private and public sector is a serious cause for concern⁷. The fallout from the algorithms used by large digital platforms, like Facebook⁸ and Twitter⁹, indeed underscores this point. The perpetuation and amplification of hate speech and misinformation, for example, through the algorithms of digital platforms is well documented.¹⁰ These acts come to engender grave offline consequences too including violence¹¹, the destabilization of democratic regimes¹² and the rapid subversion of human rights regimes¹³ that were the product of centuries of incremental progress.

The use of AI systems on a large-scale across digital platforms, notably, is not inherently problematic. These developments are transpiring rather because the systems are black boxes operating without proper oversight and regulation. These AI systems, to a great extent, shape pertinent public discussions, perpetuate certain narratives and even establish influential actors¹⁴ solely based on the virtue of their decision-making capabilities spanning across networks encompassing billions of users.

⁵ <https://www.forbes.com/sites/bernardmarr/2019/12/16/the-10-best-examples-of-how-ai-is-already-used-in-our-everyday-life/?sh=292f408d1171>

⁶ According to Stanford University's AI Index 2021 Report: <https://aiindex.stanford.edu/report/>

⁷ <https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/>

⁸ <https://www.wired.com/story/facebooks-deceptive-math-when-it-comes-to-hate-speech/>

⁹ <https://www.theverge.com/2021/10/22/22740703/twitter-algorithm-right-wing-amplification-study>

¹⁰ <https://time.com/5739688/facebook-hate-speech-languages/>

¹¹ <https://www.washingtonpost.com/technology/2021/10/22/jan-6-capitol-riot-facebook/>

¹² <https://muse.jhu.edu/article/713720>

¹³ <https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>

¹⁴ <https://ndsmcobserver.com/2021/03/changing-your-face-for-your-following-the-dark-implications-of-the-tiktok-beauty-algorithm/>

Given the proven serious implications and the fact that self-regulation of the tech industry has failed¹⁵, a conclusion reached on both sides of the Atlantic¹⁶, democratic societies can no longer rely on naïve assumptions that democracies will somehow adjust to the “new normal” and survive nor on the fact that more responsible use of AI will emerge without increased transparency and oversight. The matter is particularly pressing given that tech giants, like Metaverse, are currently devising ambitious futuristic visions that envision technology further blurring the line between online and offline spaces through the integration of social environments, office work and entertainment into one virtual reality.¹⁷

Without human rights and democracy-centered regulatory regimes for AI systems, democratic societies, institutions and individual rights will continue to erode.¹⁸ It is encouraging, however, that on both sides of the Atlantic, the need for a tiered approach to AI systems is recognized; on the EU side within the AI Act¹⁹ and on the US side via the AI Risk Management Network.²⁰ The AI Risk Management Network, nevertheless, is not a binding framework and the EU AI Act refrains from identifying AI systems used by very large online digital platforms as high risk, rendering both initiatives rather weak instruments for addressing the above-mentioned existential harms.

Recommendations

The tiered approach to regulating the use of AI in both the public and private sector by developing categories of AI systems based on the risk they pose to society should be pursued and prioritized within the Trade and Technology Council. If this regulatory approach is well defined (but also flexible), it has the potential to govern the space for the foreseeable future. This structuring would overcome the tendency for novel AI systems to outpace regulatory processes.

The emphasis on a tiered approach, therefore, should comprise a detailed and comprehensive definition of risks allowing the inclusion of new AI systems into the categories. The definitions should put a particular emphasis on ethics and human rights (broadly defined), including freedom of speech, the right to privacy and non-discrimination, and heed the protection of democracy as a core principle. Any AI system that could potentially impinge on human rights and/or be exploited to undermine democracy should be deemed high-risk and be subjected to stringent oversight and regulation.

The algorithms employed by the major online platforms²¹ that undergird the attention economy²², therefore, also need to be recognized as high risk across the transatlantic region. The platforms have experienced numerous content recommendation problems, human rights violations and misuse for information manipulation purposes and the undermining of democracy. They also could

¹⁵ <https://www.bbc.com/news/technology-51518773>

¹⁶ <https://www.forbes.com/sites/enriquedans/2021/06/12/congress-rolls-out-some-tough-regulatory-proposals-for-big-tech/?sh=27f9118c43c4>

¹⁷ <https://www.theverge.com/22588022/mark-zuckerberg-facebook-ceo-metaverse-interview>

¹⁸ <https://time.com/5930790/shoshana-zuboff-interview/>

¹⁹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>

²⁰ <https://www.nist.gov/itl/ai-risk-management-framework>

²¹ Currently defined in the EU proposed Digital Services Act as a platform with **average monthly active users in the EU** constituting at least 10% of all consumers in the EU market, <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>

²² This includes the selection of content based on the user data collection, recommendation systems, etc. https://kinit.sk/wp-content/uploads/2021/09/KINIT_Stance-of-AIA_Paper_2021_09.pdf

potentially affect substantial numbers of people within short time frames. Providing these platforms free rein without constraints or oversight is too high a cost for both individuals and societies alike to withstand.

Many critics, wary about limits placed on innovation, argue against the regulation of these new technologies. The deployment of stringent measures, however, targeted specifically towards addressing human rights violations and democratic destabilization can, on the contrary, foster innovation. It is innovative approaches, in fact, that are necessary to promote user engagement satisfying both the ethical and human rights-centered requirements imposed by democratic governments and business needs. The development and implementation of these systems by large multinational actors backed by sufficient resources and capacities (defined, for example, as very large online platforms or VLOPs in the Digital Services Act currently being debated)²³ to influence the behavior of users across the Atlantic is a prerequisite if the systems are to become adopted by smaller actors lacking similar resources.

Conclusion

Transatlantic perspectives on internet governance are urgently needed in a global environment characterized by a crisis of democracy. The rise of China, furthermore, has seen Beijing implement a dystopian online environment that employs severe censorship and social scoring systems, all presented by the country as a viable alternative to messy democratic internet governance. TTC's agenda focused on AI and a promising tiered approach to regulation is, against this backdrop, a welcome step. Both sides of the Atlantic, however, must take bold steps if the approach is to be truly efficient in safeguarding human rights and democracy. This concerted action includes the need to recognize the algorithms deployed by the digital platforms as high-risk based on their unprecedented ability to shape individual preferences and entire societies. Without a robust approach, democracies across the transatlantic space will be further undermined by opaque black boxes whose operation remains shrouded in a veil of secrecy. This status quo, if maintained, will gradually chip away at human rights and personal freedoms in a manner that could bring the dystopian alternative too close for comfort.

²³ [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/689357/EPRS_BRI\(2021\)689357_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/689357/EPRS_BRI(2021)689357_EN.pdf)