ALGORITHM WATCH / CH

ALGORITHM WATCH

/ Position by AlgorithmWatch

Input to the High Commissioner report on the practical application of the United Nations Guiding Principles on Business and Human Rights to the activities of technology companies

February 2022

Setting

Given our domain of expertise, we will focus our response to the High Commissioner's call for input on technology companies developing systems that are known as automated decision-making (ADM) systems or "Artificial Intelligence" (AI). Such systems range from algorithmic recommendation of content on platforms like Facebook/Instagram, Google/Youtube, Tiktok and others to systems used for scoring people for their creditworthiness and applications purporting to be able to derive human traits from analysing biometric features, to predict recidivism of parolees or to detect social benefit fraud. Thousands of technology companies worldwide, from global giants like Alibaba, Alphabet, Amazon, Apple, Meta, Microsoft, and Tencent to a plethora of small and medium-sized companies[1] are developing socio-technical algorithmic systems that take and prepare decisions that have far-reaching consequences for individuals, but also larger communities and, indeed, states.

[1] See e.g. the case of Clearview AI: https://news.bloomberglaw.com/privacy-and-data-security/clearview-ai-data-processing-violates-gdpr-german-regulator-says, https://www.priv.gc.ca/en/opc-news/news-and-announcements/2021/nr-c_210203/, https://techcrunch.com/2021/11/29/clearview-ai-ico-view-to-fine/

The High Commissioner will be familiar with examples of the misuse of said systems by companies like Facebook, Google, Amazon, Palantir and other technology companies. A fact that deserves as much attention is that private companies are the technology providers behind what the UN's Special Rapporteur on extreme poverty called a "digital welfare state" that uses systems "driven by digital data and technologies" to "automate, predict, identify, surveil, detect, target and punish."[2] Thus, in their delivery of social services, public authorities increasingly rely on technological systems developed by private actors.

AlgorithmWatch is a non-profit research and advocacy organization that is committed to watch, unpack and analyze automated decision-making (ADM) systems and their impact on society. While the prudent use of ADM systems can benefit individuals and communities, they come with great risks to both individual autonomy and freedom, as well as to society and the common good as a whole. Therefore, guided by the principles of autonomy, justice, harm prevention, and beneficence, and oriented at the cornerstones of democratic societies and the protection of fundamental rights, as well as the guiding principles on business and human rights, we consider it crucial to make ADM systems transparent and hold them accountable to democratic control.

We therefore welcome the opportunity to comment on the practical application of the United Nations Guiding Principles on Business and Human Rights to the activities of technology companies, as technology companies play a crucial role in the development, implementation, and use of ADM systems that affect fundamental rights. Our contribution is based on our extensive work, analysis, research and exchange with key stakeholders on the issue.[3]

We are especially concerned with a lack of transparency of ADM systems vis-à-vis not only those affected by their output, but also the greater public. This has a range of implications for the enjoyment of fundamental rights: The inaccessibility of relevant information on automated decision-making procedures impedes individuals' access to legal remedies (Art. 2(3) ICCPR), and the general opacity surrounding their use hinders public interest research, undermining foundational democratic principles based in fundamental rights. These serious risks to fundamental rights posed by ADM systems need to be addressed by a comprehensive governance framework. The following

---

[2] https://www.ohchr.org/EN/Issues/Poverty/Pages/DigitalTechnology.aspx; see also AlgorithmWatch's detailed submission https://algorithmwatch.org/en/submission-to-the-report-of-the-united-nations-special-rapporteur-on-extreme-poverty-and-human-rights/

[3] See our Automating Society Reports for the most comprehensive mapping of the use of ADM/AI-based systems in Europe: https://automatingsociety.algorithmwatch.org/

document raises key areas of concern in this regard – specifically relating to technology companies as private entities acting on behalf of public authorities as well as in their role in the private sphere. Thus, it concerns both the state's *obligation to protect* people from rights violations by third parties, as well as emerging direct human rights obligations to respect human rights of private actors, as discussed in the debate on Business and Human Rights.

Requirements on ADM systems developed or used by private actors on behalf of public authorities or in the context of PPP

Several risks to human rights can be associated with ADM systems, such as discriminatory effects (Art. 2 ICCPR, Art. 2 ICESCR), violations of the rights to freedom of expression (Art. 19 ICCPR), association (Art. 22 ICCPR) or religion (Art. 18 ICCPR), violations of the rights to freedom of assembly (Art. 21 ICCPR) and association (Art. 22 ICCPR), violations of the right to privacy (Art. 17 ICCPR) or the rights of individuals to equal treatment by the law (Art. 14 ICCPR). The use of ADM systems can also come with chilling effects on the enjoyment of these and other human rights.

These issues must be addressed comprehensively. Discriminatory, unfair, and unjust effects do not always stem from bias in training data, and thus cannot merely be avoided by ensuring data quality. ADM systems are socio-technical systems that are deployed in a certain societal context, the norms, values, and structural inequalities of which will inevitably influence the system's implications. Also, the enormous resource consumption that comes with the development and operation of ADM systems must also be taken into consideration. Regulatory frameworks therefore should address discriminatory and unjust effects of ADM systems and include explicit sustainability-related requirements.[4]

Against this background, when ADM systems are developed or used by private actors acting on behalf of public authorities or in the framework of Public-Private-Partnerships, those private actors should have to adhere to specific comprehensive safeguards. Pursuant to their obligation to protect, public authorities are required to install and enforce these safeguards.

[4] Draft AI Act: EU needs to live up to its own ambitions in terms of governance and enforcement, AlgorithmWatch, August 2021, https://algorithmwatch.org/en/eu-ai-act-consultation-submission-2021/

Case-by-case risk assessment for ADM systems

Given the unique context in which public authorities and private actors on their behalf act, the use of ADM systems must be accompanied by a systematic assessment of potential ethical implications and risks, ensuring transparency and accountability vis-à-vis those affected. These risks cannot be determined in a generalized manner, but only through a case-by-case analysis. Thus, it should be mandatory for public authorities to conduct an impact assessment prior to and during the deployment of any ADM systems, be it by themselves or by private actors on their behalf. In the latter case, public authorities shall not be exempt from the requirement to conduct such an impact assessment.

A two-stage impact assessment procedure developed by AlgorithmWatch provides a practicable and ready-to-use tool to generate transparency on such potential risks, based on seven underlying ethical principles. It enables a triage of ADM systems, indicating whether a specific system must be subject to additional transparency requirements. If this is the case, public authorities must ensure that a comprehensive transparency report is provided, allowing for the evaluation of the system and its deployment over its entire life cycle. Transparency does not by itself ensure conformity with ethical requirements, but it is a necessary condition for achieving such conformity.[5]

That is why comprehensive conformity assessments are crucial. They should not be rendered void, should not be allowed to be circumvented by outsourcing the development and deployment of ADM systems to private actors, should take place within an unambiguous governance structure, and adequately include third-party oversight. These are essential steps in ensuring safeguards from potential risks associated with ADM systems throughout their entire life cycle.[6]

Public Registers

AlgorithmWatch is calling for a publicly accessible list of (i) all ADM systems deployed in the public sector, regardless of whether deployed by public authorities or by private actors acting on their behalf, and (ii) ADM systems operated by private actors, whenever their use has a significant impact on an individual, a specific group, or society at large.

---

[5] Automated Decision-Making Systems in the Public Sector – An Impact Assessment Tool for Public Authorities, Michele Loi in collaboration with Anna Mätzener, Angela Müller, and Matthias Spielkamp, June 2021, https://algorithmwatch.org/en/adms-impact-assessment-public-sector-algorithmwatch/

[6] Draft AI Act: EU needs to live up to its own ambitions in terms of governance and enforcement, AlgorithmWatch, August 2021, https://algorithmwatch.org/en/eu-ai-act-consultation-submission-2021/

Next to information on its purpose, the underlying model, and the developers and deployers of the system, this register should contain the results of any impact assessment undertaken.

In cases where full public disclosure cannot be granted – for legitimate reasons that need to be clearly and narrowly defined – the database should include information of the respective body to which such full transparency was granted (for example, the national supervisory authority). Further, the information must be provided in an easily readable and accessible manner, including structured digital data based on a standardized protocol.

Accountability

Moreover, transparency – as a mere instrumental means to an end – must be coupled with adequate and reliable accountability frameworks. The bearers of fundamental rights must be equipped with the means to defend themselves if they feel they have been treated unlawfully (Art. 2(3), 14 ICCPR). When ADM systems have consequential effects on people's lives, then people must be able to retrieve all relevant information about what has led to the outcome of the corresponding decision – regardless of whether the system was operated by a public or a private actor. In the public sector, individuals who were subject to a decision that is solely or significantly informed by the output of an ADM system should be notified without delay.

Inter alia, holding ADM systems accountable also requires making training data and data results accessible to independent researchers, journalists, and civil society organizations. We thus suggest introducing legally binding data access frameworks, focused explicitly on supporting and enabling public interest research and in full respect of data protection and privacy law. This contributes to shedding light on the so far opaque use of ADM systems and to enabling an urgently needed evidence-based debate on the topic.

Those affected must also have the possibility to legally challenge outcomes of ADM systems used by both public and private actors. There must be easily accessible, affordable, and effective legal remedies at hand for affected individuals and groups to contest such decisions and, where appropriate, to demand reversal, reconsideration through a different procedure, or compensation. Mere technological solutions do not suffice to ensure that AI systems are used to the benefit of the many, not the few.

We expect policymakers to step up for the bearers of fundamental rights and to introduce effective accountability mechanisms for those affected by systems' outputs, ensuring that those with less bargaining power have a voice, too. This includes a right to information, data access frameworks for public interest research, and effective legal remedies.

Limited liability for online content moderation

A substantial part of public discourse today takes place on digital media platforms, which are largely governed by private actors. As a consequence, substantial questions on, for example the scope of the right to freedom of opinion and expression (Art. 19 ICCPR) are opaquely decided by large online platforms – questions that should, from a rule of law perspective, be addressed by the judiciary in the framework of careful and transparent legal analyses. Also, public authorities bound by fundamental rights cannot ban "low-quality" content or demand its suppression as long as it is legal.[7] This is a fundamental basis of freedom of expression and must not be undermined. That is why a limited liability framework, such as the one outlined in the European Union's E-Commerce directive and updated in its proposed Digital Services Act, is the right general approach to dealing with illegal user-generated content. This framework includes a ban on a general monitoring obligation (also known as upload filters) which could lead to the overblocking of legal speech, as well as rules that empower users to settle disputes on the legality of content through independent "dispute settlement bodies."

A successful limited liability regime needs to be complemented by transparency frameworks and systems to effectively audit algorithmic systems that empower independent third parties to hold platforms to account and thereby help to, at least partly, relieve the individual user from the burden of proof.

[7] Matthias Cornils et al (2020): Designing Platform Governance: A Normative Perspective on Regulatory Needs, Strategies, and Tools to Enhance the Information Function of Intermediaries; at the same time international human rights law (e.g. Art. 15 ECHR) [https://algorithmwatch.org/en/wp-content/uploads/2020/10/Governing-Platforms_DSA-Recommendations.pdf], puts very strict requirements for the conditions under which states can restrict freedom of expression and information, notably the principles of legality, necessity and proportionality and legitimacy.

Transparency and accountability for large platforms

Due to the crucial role that large tech platforms play in society, influencing vital interactions from identity building to voting choices, we need more transparency to ensure an evidence-based debate on their impact – which is a necessary step towards holding them accountable. Existing transparency tools have failed to provide watchdogs and regulators with the information they need to hold large platforms accountable for their impact on democratic processes and fundamental rights.[8]

One of the key barriers to ensuring adequate transparency of algorithmic systems is the lack of reliable access to the data that watchdogs need to scrutinize how very large platforms target, moderate, and recommend content or services to their users, as well as a hostile approach from key platforms toward public interest scrutiny.[9] This is why we believe in the need for comprehensive data access frameworks for public interest research. Since voluntary data access frameworks by platforms have proven to be unreliable or even useless, such data access frameworks must be legally introduced and publicly managed or mandated to an intermediate institution. Only if we understand how our public sphere is influenced by platforms' algorithmic choices can we take measures towards ensuring they do not undermine individuals' autonomy, freedom, and the collective good.

Red Line on Biometric Mass Surveillance

Some uses of AI-systems are inherently incompatible with fundamental rights. Biometric recognition systems in public spaces enable undifferentiated mass surveillance, which inherently conflicts with fundamental rights. If people can be identified and monitored in public space at any time, this not only violates their right to privacy but also has a chilling effect that prevents them from exercising fundamental rights such as those to freedom of expression or assembly (Art. 19, 21 ICCPR) – freedoms that are essential for participation in public life and public discourse in a democratically organized society. For already disadvantaged groups or minorities, these effects typically manifest themselves in amplified form.

---

[8] Are Algorithms a Threat to Democracy? The Rise of Intermediaries: A Challenge for Public Discourse, Professor Dr. Birgit Stark and Daniel Stegmann, M.A. with Melanie Magin, Assoc. Prof. & Dr. Pascal Jürgens, May 2020, https://algorithmwatch.org/en/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch.pdf

[9] Under Facebook's thumb: Platforms must stop suppressing public interest research, AlgorithmWatch, August 2021, https://algorithmwatch.org/en/defend-public-interest-research-on-platforms/

Biometric mass surveillance not only represents an encroachment on the fundamental rights of those being monitored, but also damages democracy as a whole. AlgorithmWatch therefore calls for an explicit ban on the use of biometric mass surveillance technologies in publicly accessible spaces – thus addressing both public and private actors developing and deploying such technologies.[10]

Labor rights

Complex computational systems are increasingly being used to monitor, score, manage, promote, and even fire employees. These systems have the potential to profoundly influence, alter, and redirect the lives of people at work and, therefore, impact their life opportunities in general. The use of these systems is often non-transparent and happens without prior consent, can undermine autonomy and participation, and risks turning employees into mere objects (violating Art. 1 UDHR). In order to protect workers' rights – and, first and foremost, their autonomy (Art. 7 ICESCR) – there must be mandatory provisions for transparency and participation that technology companies developing such systems and employers adhere to. States are obliged to protect people by taking the necessary measures with respect to the use of AI-based systems in the workplace.

Workers must be guaranteed the right to obtain information about a system's purpose and how that purpose is intended to be achieved. This does not necessarily imply full disclosure of an algorithm, a model, source code, or data. An appropriate level of intelligibility can be achieved by providing information about key qualities of a system, e.g., using guidelines for reviewing essential features. If such a level of understanding cannot be sufficiently achieved by these means, then there needs to be a procedure defined to compel providers and users to make satisfactory information available.

Company stakeholders and worker representatives need to be supported in effective capacity building in relation to AI-based systems. State actors should either provide training courses themselves, or allocate resources for third parties to do so. Also, anti-discrimination laws should be evaluated in line with technological developments. As we have shown, online platforms optimize advertisement delivery in discriminatory ways that very much affect the choices of jobseekers. These practices are often not adequately addressed by current anti discrimination laws.

---

[10] Open letter calling for a global ban on biometric recognition technologies that enable mass and discriminatory surveillance, AlgorithmWatch, June 2021, https://algorithmwatch.org/en/open-letter-ban-biometric-surveillance/

Enforcement

To make a difference on the ground, policy-makers should take into account enforcement procedures when laying out regulatory approaches.[11] National authorities must be equipped with the expertise and resources to effectively fulfill their tasks. They must be sufficiently independent, adequately resourced, and have the relevant expertise – in both technology and fundamental rights – to fulfill the tasks assigned to them.

Final remarks

Two major European legislative proposals, the AI Act and Digital Services Act, include a wide range of policy responses with implications for how ADM systems affect human rights. AlgorithmWatch has published extensive analysis of both of these proposals and offered policy recommendations aimed at ensuring the rules governing ADM systems in the European Union are aligned with fundamental democratic and human rights. Further, AlgorithmWatch, in its role as an observer, has extensively contributed to the work of the Council of Europe's Ad Hoc Committee on Artificial Intelligence (CAHAI), in examining the feasibility and possible elements of a legal framework for the development, design and application of AI systems, based on the Council of Europe's standards on human rights, democracy and the rule of law. We will continue to do so during the negotiations on a legal framework on AI within the Council of Europe.

AlgorithmWatch is looking forward to further contributing to the debate within the UN, given the importance of multilateral governance frameworks in this field.

[11] Flaws in ex-post enforcement in the AI Act, Irish Council for Civil Liberties, 15 February 2022, https://www.iccl.ie/wp-content/uploads/2022/02/20220215_ICCL_AIActEnforcementLetter.pdf