/ Contribution by AlgorithmWatch

Input to Report on "The Right to Privacy in the Digital Age (2022)"

Focus of our contribution

The present contribution will focus on **algorithmic systems,** often referred to under the term **'Artificial Intelligence' (AI).** The use of algorithmic systems can **impact the enjoyment of human rights in a variety of ways.** This contribution will not only consider the right to privacy but also related human rights norms, as asked for in the call for inputs, focusing on the International Covenant on Civil and Political Rights (ICCPR).

Overview over main challenges and potential negative impact on human rights

First, algorithmic systems are often deployed as «black boxes», thus in a highly opaque manner. This **lack of transparency** vis-à-vis not only those affected by their output but also the greater public has a range of implications for the enjoyment of human rights: The inaccessibility of relevant information on the algorithmic decision-making procedures impedes individuals' access to legal remedies, and the general opacity on their use hinders academics, civil society, and journalists from conducting public interest research, undermining foundational democratic principles based in human rights (Art. 2(3), 14, 15, 16, 19, 25(c) ICCPR).

Second, and as widely debated, algorithmic systems can have **discriminatory effects** (Art. 26 ICCPR). For example, they can adopt, perpetuate, and even reinforce the bias existing in their training data, if this data is not representative of the group affected by the system—and because existing data always mirrors the biases and discriminatory tendencies of our society.

However, third, even if the problem of bias in training data was to be solved at the technological level and accuracy of the system improved, this does not avoid the potentially detrimental impact on human rights and the unjust impacts such systems can have. For example, algorithmic systems used in the public sector can still lead to **discrimination** and **unequal treatment** (Art. 3, 25(c), 26 ICCPR), insofar as they might disadvantage certain groups, typically those already in a vulnerable position—e.g., those receiving social benefits or asylum seekers. Thus, even if 'debiasing' a system worked at the technological level, detrimental impacts on human rights can come about by the context and the way in which a system is put to use.

Fourth, some uses of algorithmic systems are **inherently incompatible with human rights**. E.g., the use of biometric recognition systems in public space can enable forms of mass surveillance that can never be conducted in compliance with human rights. As such, it violates people's rights to privacy (Art. 17 ICCPR), and it creates a chilling effect that undermines rights to freedom of thought, conscience and religion (Art. 18 ICCPR), freedom of opinion, expression and information (Art. 19 ICCPR), and freedom of assembly and association (Art. 21, 22 ICCPR).

Fifth, a substantial part of public discourse today takes place in the digital sphere, which is largely governed by private actors, creating a protection vacuum: Human rights obligations do not directly apply to private actors. Thus, they are not directly obliged to respect rights that are easily threatened in this digital public sphere, such as the right to freedom of opinion, expression and information (Art. 19 ICCPR). Furthermore, today, substantial questions on, e.g., the scope of freedom of expression 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.

Sixth, as to the digital transition of work, the increasing use of algorithmic systems in recruiting and as 'people analytics' tools threatens human rights of affected individuals at their workplace. Their use is often non-transparent and happens without prior consent, can undermine autonomy and participation, and risks turning employees into mere objects, undermining their inherent dignity (Preamble ICCPR). As private employers are not subject to direct human rights obligations, states are obliged to protect people by taking the necessary measures with respect to the use of algorithmic systems in the workplace. Vis-à-vis their own personnel, they are required to use algorithmic systems in ways that comply to human rights standards.

Lastly, the massive computational infrastructure projects implemented in response to the Covid-19-pandemic, which were often realized in the framework of public-private partnerships, illustrate the need to strengthen such governance. In general, the public may lose control—and thus autonomy—over its decision-making processes if it relies on infrastructure entirely owned and made (in)accessible by third parties. The lifespan of digital tools used in the fight against the pandemic should be clearly limited in order to avoid that they are used for other than their original purposes, which may interfere with a range of fundamental rights (including Art. 12, 17, 18, 19, 21, 26 ICCPR).

Examples of issues and situations in which violations are likely to occur

Biometric recognition systems in public space

The use of biometric recognition systems in public space not only threatens people's right to privacy (Art. 17 ICCPR), but also creates a chilling effect that restricts their enjoyment of the rights to freedom of thought, conscience and religion (Art. 18 ICCPR), freedom of opinion, expression and information (Art. 19 ICCPR), and freedom of assembly and of association (Art. 21, 22 ICCPR).

Example Italy: Italy's Data Protection Authority blocked SARI Real Time facial recognition system acquired by the Police on grounds of missing legal basis. The DPA also argued that future attempts to install such systems would need to be in compliance with legal regulations.

Use of algorithmic systems on people in vulnerable positions (migrants, inmates)

The use of algorithmic systems on marginalized communities, such as migrants and asylum seekers, can reinforce the vulnerable position they find themselves in. The use of these tools can be of an intrusive nature, and it can multiply the sometimes already non-transparent nature of public services, e.g., the processing of asylum requests, which makes it more difficult for those affected to access relevant information and to gain necessary knowledge in order to legally contest decisions (Art. 2(3), 14, 15, 16, 17, 18, 19, 26; Art. 14 UDHR).

Example Germany: The Federal Office for Migration and Refugees (BAMF) has been using automated text and speech recognition systems to identify refugees (Thüer, Köver and Fanta, 2018) since 2017. Agency employees can ask asylum seekers to give them access to their cell phone, tablet, or laptop to verify if they are telling the truth about where they come from, can obtain all the data contained on them and run software on it. The software presents the employee with a limited overview of the content, which also includes language analysis of the text retrieved. Software and hardware are provided by private companies (Biselli, 2017; Biselli, 2018b). Another tool deployed by the BAMF aims to identify disguised dialects in speech (Biselli, 2018a). When an asylum seeker does not have a valid proof of ID, a two-minute voice recording of the person describing a picture in their mother tongue is analyzed by software, calculating a percentage of how close the speech comes to a certain dialect.

Example Spain: Eleven years ago, the Catalonian Department of Justice, in Spain, introduced RisCanvi, a system that estimates the risk that inmates reoffend upon leaving prison. All Catalonian prisons use it, but the tool is hardly transparent. RisCanvi ("risk change" in Catalan) does not decide whether or not an inmate will be paroled. The final decision is made by professionals of the justice system, who do not necessarily have to agree with the result of the algorithm's analysis. Nevertheless, it seems that in the great majority of the cases, the evaluation of the professionals follows the algorithm's.

Broad trends with regard to the impact of algorithmic systems on human rights

There are three broad trends we regard as among the most pertinent threats to human rights by algorithmic systems.

In the years to come, the **automation of procedures and services in public administrations** is likely to increase exponentially. Citizens demand user-friendly services available 24/7, administrations regard automation as a chance to accelerate efficiency, facilitate processes, and expedite mass and routine services. However, given the unique context in which public authorities act, the deployment of algorithmic systems should be accompanied by a systematic evaluation of potential ethical implications, ensuring transparency and accountability vis-à-vis those affected.

While risks are not limited to the public sector (but often mirror similar risks pertaining to the private sphere), the latter is of a special kind. Public authorities are not only subject to unique legal preconditions, such as to the principles of legality or compliance with human rights. They also act in a unique setting where individuals do not have the freedom to choose the provider of services but are inescapably subject to a particular administration, according to rules of jurisdictional authority. In addition, decisions by public authorities often have consequential effects on individuals. When deploying algorithmic systems in the public sector, this context must be considered—individual and societal trust should be the ultimate benchmarks for the automation of administrative procedures.

A further area of concerns is the **automation of the work sector**, which must be accompanied by reliable measures at domestic levels, ensuring and defending employee's rights to transparency and to fair working conditions, protecting their autonomy and dignity at the workplace.

Lastly, the **digital nature of the public sphere** is likely to further increase—a development that during the report period was boosted by the pandemic. In order to safeguard a healthy and resilient public sphere, light must be shed on the functioning and role of online platforms and intermediaries, their impact on democratic societies be mitigated, and frameworks to guarantee accountability vis-à-vis those affected must be set up or strengthened. It is crucial to stress the inherent link between human rights and democracy: Human rights are a necessary precondition of democracy—and for democratic systems to work. Accordingly, impacts of algorithmic systems on individuals and society cannot be regarded as two separate topics but must always be comprehensively analyzed and addressed.