Advance Edited Version

Distr.: General
13 September 2021

Original: English

Human Rights Council
Forty-eighth session
13 September–1 October 2021
Agenda items 2 and 3
Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General

Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development

The right to privacy in the digital age*

Report of the United Nations High Commissioner for Human Rights

Summary

In the present report, mandated by the Human Rights Council in its resolution 42/15, the High Commissioner analyses how the widespread use by States and businesses of artificial intelligence, including profiling, automated decision-making and machine-learning technologies, affects the enjoyment of the right to privacy and associated rights. Following an overview of the international legal framework, the High Commissioner highlights aspects of artificial intelligence that facilitate interference in privacy and provides examples of impacts on the right to privacy and associated rights in four key sectors. The High Commissioner then discusses approaches to addressing the challenges, providing a set of recommendations for States and businesses regarding the design and implementation of safeguards to prevent and minimize harmful outcomes and to facilitate the full enjoyment of the benefits that artificial intelligence can provide.

* The present report was submitted after the deadline so as to include the most recent information.

I. Introduction

1. The present report is submitted pursuant to Human Rights Council resolution 42/15, in which the Council requested the United Nations High Commissioner for Human Rights to organize an expert seminar to discuss how artificial intelligence, including profiling, automated decision-making and machine-learning technologies may, without proper safeguards, affect the enjoyment of the right to privacy, to prepare a thematic report on the issue and to submit it to the Council at its forty-fifth session.[1]

2. No other technological development of recent years has captured the public imagination more than artificial intelligence (AI), in particular machine-learning technologies.[2] Indeed, these technologies can be a tremendous force for good, helping societies overcome some of the great challenges of the current time. However, these technologies can also have negative, even catastrophic, effects if deployed without sufficient regard to their impact on human rights.

3. While the present report does not focus on the coronavirus disease (COVID-19) pandemic, the ongoing global health crisis provides a powerful and highly visible example of the speed, scale and impact of AI in diverse spheres of life across the globe. Contact-tracing systems using multiple types of data (geolocation, credit card, transport system, health and demographic) and information about personal networks have been used to track the spread of the disease. AI systems have been used to flag individuals as potentially infected or infectious, requiring them to isolate or to quarantine. AI systems used for the predictive allocation of grades resulted in outcomes that discriminated against students from public schools and poorer neighbourhoods. These developments have demonstrated the broad range of impacts that AI systems have on people's daily lives. The right to privacy is affected in all these cases, with AI using personal information and often making decisions that have tangible effects on people's lives. Nevertheless, deeply intertwined with the question of privacy are various impacts on the enjoyment of other rights, such as the rights to health, education, freedom of movement, freedom of peaceful assembly, freedom of association and freedom of expression.

4. In 2019, in "The highest aspiration: a call to action for human rights", the Secretary-General of the United Nations recognized that the digital age had opened up new frontiers of human welfare, knowledge and exploration. He underscored that digital technologies provide new means to advocate for, defend and exercise human rights. Nevertheless, new technologies are too often used to violate rights, especially those of people who are already vulnerable or being left behind, for instance through surveillance, repression, censorship and online harassment, including of human rights defenders. The digitization of welfare systems, despite its potential to improve efficiency, risks excluding the people who are most in need. The Secretary-General emphasized that advances in new technologies must not be used to erode human rights, deepen inequality or exacerbate existing discrimination. He stressed that the governance of AI needs to ensure fairness, accountability, explainability and transparency. In the security sphere, the Secretary-General reiterated his call for a global prohibition on lethal autonomous weapon systems.

5. The present report builds upon the two previous reports of the High Commissioner on the issue of the right to privacy in the digital age.[3] It also incorporates the insights gained in the virtual expert seminar that was organized pursuant to Council resolution 42/15, held on 27 and 28 May 2020, as well as the responses to the High Commissioner's call for input to the present report.[4]

[1] The preparation of the report was postponed. See A/HRC/45/26 and A/HRC/47/61.

[2] There is no generally accepted definition of the term "artificial intelligence". In the present report, it is employed to refer to a constellation of processes and technologies enabling computers to complement or replace specific tasks otherwise performed by humans, such as making decisions and solving problems (A/73/348, para. 3), which includes but is not limited to machine learning and deep learning.

[3] A/HRC/27/37 and A/HRC/39/29.

[4] See www.ohchr.org/EN/Issues/DigitalAge/Pages/cfi-digital-age.aspx for the call for input and the submissions received.

II. Legal framework

6. Article 12 of the Universal Declaration of Human Rights, article 17 of the International Covenant on Civil and Political Rights and several other international and regional human rights instruments recognize the right to privacy as a fundamental human right.[5] The right to privacy plays a pivotal role in the balance of power between the State and the individual and is a foundational right for a democratic society.[6] Its importance for the enjoyment and exercise of other human rights online and offline[7] in an increasingly data-centric world is growing.

7. The right to privacy is an expression of human dignity and is linked to the protection of human autonomy and personal identity.[8] Aspects of privacy that are of particular importance in the context of the use of AI include informational privacy, covering information that exists or can be derived about a person and her or his life and the decisions based on that information,[9] and the freedom to make decisions about one's identity.

8. Any interference with the right to privacy must not be arbitrary or unlawful.[10] The term "unlawful" means that States may interfere with the right to privacy only on the basis of law and in accordance with that law. The law itself must comply with the provisions, aims and objectives of the International Covenant on Civil and Political Rights and must specify in detail the precise circumstances in which such interference is permissible.[11] The introduction of the concept of arbitrariness is intended to guarantee that even interference provided for by law should be in accordance with the provisions, aims and objectives of the Covenant and should, in any event, be reasonable in the particular circumstances.[12] Accordingly, any interference with the right to privacy must serve a legitimate purpose, be necessary for achieving that legitimate purpose and be proportionate.[13] Any restriction must also be the least intrusive option available and must not impair the essence of the right to privacy.[14]

9. The right to privacy applies to everyone. Differences in its protection on the basis of race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status are inconsistent with the principle of non-discrimination laid down in articles 2 (1) and 3 of the International Covenant on Civil and Political Rights. Discrimination on these grounds also violates the right to equality before the law contained in article 26 of the Covenant.

---

[5] See article 16 of the Convention on the Rights of the Child, article 14 of the International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, article 22 of the Convention on the Rights of Persons with Disabilities, article 10 of the African Charter on the Rights and Welfare of the Child, article 11 of the American Convention on Human Rights and article 8 of the Convention for the Protection of Human Rights and Fundamental Freedoms (the European Convention on Human Rights).

[6] A/HRC/39/29, para. 11.

[7] Committee on the Rights of the Child, general comment No. 25 (2021), paras. 67–68; and A/HRC/39/29, para. 11.

[8] Committee on the Rights of the Child, general comment No. 25 (2021), para. 67; and European Court of Human Rights, *Goodwin v. the United Kingdom*, Application No. 28957/95, Judgment of 11 July 2002, para. 90.

[9] A/HRC/39/29, para. 5.

[10] For a detailed analysis of the terms "arbitrary" and "unlawful", see A/HRC/27/37, paras. 21–27.

[11] Human Rights Committee, general comment No. 16 (1988), paras. 3 and 8.

[12] Ibid., para. 4.

[13] *Toonen v. Australia* (CCPR/C/50/D/488/1992), para. 8.3, *Van Hulst v. Netherlands* (CCPR/C/82/D/903/1999), paras. 7.3 and 7.6, *Madhewoo v. Mauritius* (CCPR/C/131/D/3163/2018), para. 7.5, and CCPR/C/USA/CO/4, para. 22. See also Committee on the Rights of the Child, general comment No. 25 (2021), para. 69.

[14] Human Rights Committee, general comment No. 31 (2004), para. 6; A/HRC/27/37, para. 22, and A/HRC/39/29, para. 10.

10. Article 2 (1) of the International Covenant on Civil and Political Rights requires States to respect and ensure the rights recognized in the Covenant for all individuals within their territory and subject to their jurisdiction, without discrimination. In other words, States must not only refrain from violating the rights recognized in the Covenant,[15] but they also have an obligation to take positive steps to protect the enjoyment of those rights. This implies a duty to adopt adequate legislative and other measures to safeguard individuals against interference in their privacy, whether it emanates from State authorities or from natural or legal persons.[16] This duty is also reflected in pillar I of the Guiding Principles on Business and Human Rights, which outlines the duty of States to protect against adverse human rights impacts involving companies.

11. Business enterprises have a responsibility to respect all internationally recognized human rights. This means that they should avoid infringing on the human rights of others and address adverse human rights impacts with which they are involved. Pillar II of the Guiding Principles on Business and Human Rights provides an authoritative blueprint for all enterprises regarding how to meet this responsibility.[17] The responsibility to respect applies throughout an enterprise's activities and business relationships.

III. Impacts of artificial intelligence on the right to privacy and other human rights

A. Relevant features of artificial intelligence systems

12. The operation of AI systems can facilitate and deepen privacy intrusions and other interference with rights in a variety of ways. These include entirely new applications as well as features of AI systems that expand, intensify or incentivize interference with the right to privacy, most notably through increased collection and use of personal data.

13. AI systems typically rely on large data sets, often including personal data. This incentivizes widespread data collection, storage and processing. Many businesses optimize services to collect as much data as possible.[18] For example, online businesses like social media companies rely on the collection and monetization of massive amounts of data about Internet users.[19] The so-called Internet of Things is a rapidly growing source of data exploited by businesses and States alike. Data collection happens in intimate, private and public spaces.[20] Data brokers acquire, merge, analyse and share personal data with countless recipients. These data transactions are largely shielded from public scrutiny and only marginally inhibited by existing legal frameworks.[21] The resulting data sets are large and the information collected is of unprecedented proportions.

14. Apart from exposing people's private lives to companies and States, these data sets make individuals vulnerable in a number of other ways. Data breaches have repeatedly exposed sensitive information of millions of people.[22] Large data sets enable countless forms of analysis and sharing of data with third parties, often amounting to further privacy

[15] Human Rights Committee, general comment No. 31 (2004), para. 6.

[16] A/HRC/39/29, para. 23. See also Human Rights Committee, general comments No. 16 (1988), paras. 1 and 9, and No. 31 (2004), para. 8; and Committee on the Rights of the Child, general comment No. 25 (2021), paras. 36–39.

[17] In its resolution 17/4, the Human Rights Council unanimously endorsed the Guiding Principles on Business and Human Rights.

[18] Wolflie Christl, *Corporate surveillance in everyday life* (Vienna, Cracked Lab – Institute for Critical Digital Culture, 2017).

[19] Submission by Ranking Digital Rights.

[20] Submissions by Centre for Communication Governance at National Law University Delhi, Derechos Digitales, Digital Rights Watch, Global Partners Digital, International Center for Not-for-Profit Law and Universidade Federal de Uberlândia.

[21] Aaron Rieke and others, *Data brokers in an open society* (London, Open Society Foundation, 2016).

[22] See, e.g., www.ftc.gov/news-events/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related.

intrusions and incurring other adverse human rights impacts. Arrangements enabling government agencies to have direct access to such data sets held by businesses, for example, increase the likelihood of arbitrary or unlawful interference in the right to privacy of the individuals concerned.[23] One particular concern is the possibility of de-anonymization that is facilitated by fusing data from various sources.[24] At the same time, the design of data sets can have implications for individuals' identity. For example, a data set that records gender as binary misgenders those who do not identify as male or female. Long-term storage of personal data also carries particular risks, as data are open to future forms of exploitation not envisaged at the time of data collection.[25] Over time, the data can become inaccurate, irrelevant or carry over historic misidentification, thereby causing biased or erroneous outcomes of future data processing.[26]

15. It should be noted that AI systems do not exclusively rely on the processing of personal data. However, even when personal data are not involved, human rights, including the right to privacy, may still be adversely affected by their use,[27] as shown below.

16. AI tools are widely used to seek insights into patterns of human behaviour. With access to the right data sets, it is possible to draw conclusions about how many people in a particular neighbourhood are likely to attend a certain place of worship, what television shows they may prefer and even roughly what time they tend to wake up and go to sleep. AI tools can make far-reaching inferences about individuals, including about their mental and physical condition,[28] and can enable the identification of groups, such as people with particular political or personal leanings. AI is also used to assess the likelihood of future behaviour or events. AI-made inferences and predictions, despite their probabilistic nature, can be the basis for decisions affecting people's rights, at times in a fully automated way.

17. Many inferences and predictions deeply affect the enjoyment of the right to privacy, including people's autonomy and their right to establish details of their identity. They also raise many questions concerning other rights, such as the rights to freedom of thought and of opinion, the right to freedom of expression, and the right to a fair trial and related rights.

18. AI-based decisions are not free from error. In fact, the scalability of AI solutions can dramatically increase negative effects of seemingly small error rates.[29] Faulty outputs of AI systems have various sources. To start with, outputs of AI algorithms have probabilistic elements, which means that there is uncertainty attached to their outputs.[30] Moreover, the relevance and accuracy of data used are often questionable. Furthermore, unrealistic expectations can lead to the deployment of AI tools that are not equipped to achieve the desired goals. For example, an analysis of hundreds of medical AI tools for diagnosing and predicting COVID-19 risks, developed with high hopes, revealed that none of them had been fit for clinical use.[31]

19. Outputs from AI systems relying on faulty data can contribute to human rights violations in a multitude of ways, for example, by erroneously flagging an individual as a

[23] Submission by Global Network Initiative.

[24] Submissions by Centre for Communication Governance at National Law University Delhi, Derechos Digitales and Privacy International.

[25] Submission by OVD-Info.

[26] Committee on the Elimination of Racial Discrimination, general recommendation No. 36 (2020), para. 33.

[27] Council of Europe, "Guidelines on addressing the human rights impacts of algorithmic systems", (appendix to Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems), sect. A, para. 6.

[28] Submissions by Derechos Digitales and Privacy International.

[29] Submission by Germany.

[30] European Union Agency for Fundamental Rights, "#BigData: Discrimination in data-supported decision-making" (Vienna, 2018), p. 4.

[31] See www.technologyreview.com/2021/07/30/1030329/machine-learning-ai-failed-covid-hospital-diagnosis-pandemic/.

likely terrorist or as having committed welfare fraud. Biased data sets that lead to discriminatory decisions based on AI systems are particularly concerning.[32]

20. The decision-making processes of many AI systems are opaque. The complexity of the data environment, algorithms and models underlying the development and operation of AI systems, as well as the intentional secrecy of government and private actors are factors that undermine meaningful ways for the public to understand the effects of AI systems on human rights and society. Machine-learning systems add an important element of opacity; they can be capable of identifying patterns and developing prescriptions that are difficult or impossible to explain.[33] This is often referred to as the "black box" problem.[34] The opacity makes it challenging to meaningfully scrutinize an AI system and can be an obstacle for effective accountability in cases where AI systems cause harm.[35] Nevertheless, it is worth noting that these systems do not have to be entirely inscrutable.[36]

B. Concerns about artificial intelligence systems in key sectors

21. The present section illustrates how these concerns are experienced in practice by considering four key areas where the application of AI tools has given rise to concern.

Artificial intelligence in law enforcement, national security, criminal justice and border management

22. States are increasingly integrating AI systems into law enforcement, national security, criminal justice and border management[37] systems. While many of these applications may indeed be a cause for concern, the present section will focus on a few select examples that represent some of the diverse emerging human rights issues.

23. AI systems are often used as forecasting tools. They use algorithms to analyse large quantities of data, including historic data, to assess risks and predict future trends. Depending on the purpose, training data and data analysed can include, for example, criminal records, arrest records, crime statistics, records of police interventions in specific neighbourhoods, social media posts, communications data and travel records.[38] The technologies may be used to create profiles of people, identify places as likely to be sites of increased criminal or terrorist activity, and even flag individuals as likely suspects and future reoffenders.[39]

24. The privacy and broader human rights implications of these activities are vast. First, the data sets used include information about large numbers of individuals, thus implicating their right to privacy. Second, they can trigger interventions by the State, such as searches, questioning, arrest and prosecution, even though AI assessments by themselves should not be seen as a basis for reasonable suspicion due to the probabilistic nature of the predictions. Rights affected include the rights to privacy, to a fair trial, to freedom from arbitrary arrest and detention and the right to life. Third, the inherent opacity of AI-based decisions raises particularly pressing questions concerning State accountability when AI informs coercive

[32] Committee on the Elimination of Racial Discrimination, general recommendation No. 36 (2020), paras. 31–36; and High-level Panel on Digital Cooperation, "The age of digital interdependence" (June 2019), pp. 17–18.

[33] Submission by Germany.

[34] See www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai/.

[35] See www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6; and www.cambridge.org/core/journals/international-review-of-the-red-cross/article/abs/ai-for-humanitarian-action-human-rights-and-ethics/C91D044210CADF7A0E023862CF4EE758.

[36] See, e.g., Inioluwa Deborah Raji and others, "Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing", 3 January 2020.

[37] For an in-depth analysis of the human rights implications of AI and other digital technologies in border management, see A/75/590.

[38] Submission by Privacy International. See also A/HRC/44/57, para. 35.

[39] See www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf.

measures, even more so in areas that typically suffer from a general lack of transparency, such as the activities of counter-terrorism forces.[40] Fourth, predictive tools carry an inherent risk of perpetuating or even enhancing discrimination, reflecting embedded historic racial and ethnic bias in the data sets used, such as a disproportionate focus of policing of certain minorities.[41]

25. Developments in the field of biometric recognition technology have led to its increasing use by law enforcement and national security agencies. Biometric recognition relies on the comparison of the digital representation of certain features of an individual, such as the face, fingerprint, iris, voice or gait, with other such representations in a database.[42] From the comparison, a higher or lower probability is deduced that the person is indeed the person to be identified. These processes are increasingly carried out in real time and from a distance. In particular, remote real-time facial recognition is increasingly deployed by authorities across the globe.[43]

26. Remote real-time biometric recognition raises serious concerns under international human rights law, which the High Commissioner has highlighted previously.[44] Some of these concerns reflect the problems associated with predictive tools, including the possibility of erroneous identification of individuals and disproportionate impacts on members of certain groups.[45] Moreover, facial recognition technology can be used to profile individuals on the basis of their ethnicity, race, national origin, gender and other characteristics.[46]

27. Remote biometric recognition is linked to deep interference with the right to privacy. A person's biometric information constitutes one of the key attributes of her or his personality as it reveals unique characteristics distinguishing her or him from other persons.[47] Moreover, remote biometric recognition dramatically increases the ability of State authorities to systematically identity and track individuals in public spaces, undermining the ability of people to go about their lives unobserved and resulting in a direct negative effect on the exercise of the rights to freedom of expression, of peaceful assembly and of association, as well as freedom of movement.[48] Against this background, the High Commissioner therefore welcomes recent efforts to limit or ban the use of real-time biometric recognition technologies.[49]

28. AI tools have also been developed to allegedly deduce people's emotional and mental state from their facial expressions and other "predictive biometrics" to decide whether they are a security threat.[50] Facial emotional recognition systems operate on the premise that it is possible to automatically and systematically infer the emotional state of human beings from

---

[40] A/74/335 and A/HRC/43/46, paras. 37–38.

[41] Submission by Tech Hive Advisory Limited. See also Committee on the Elimination of Racial Discrimination, general recommendation No. 36 (2020), para. 33; and the conference room paper of the United Nations High Commissioner for Human Rights on promotion and protection of the human rights and fundamental freedoms of Africans and of people of African descent against excessive use of force and other human rights violations by law enforcement officers (A/HRC/47/CRP.1), available at www.ohchr.org/Documents/Issues/Racism/A_HRC_47_CRP_1.pdf, paras. 15 and 19.

[42] A/HRC/31/64, para. 14.

[43] Submissions by Derechos Digitales and International Center for Not-for-Profit Law.

[44] A/HRC/44/24.

[45] Submission by Privacy International.

[46] A/HRC/44/57, paras. 39–40.

[47] A/HRC/44/24, para. 33. See also European Court of Human Rights, *Reklos and Davourlis v. Greece*, Application No. 1234/05, Judgment of 15 April 2009, para. 40.

[48] See European Data Protection Board and European Data Protection Supervisor, joint opinion 5/2021, para. 30; and submissions by International Center for Not-for-Profit Law and Privacy International. See also A/HRC/44/24, para. 34, and A/HRC/41/35.

[49] Submission by European Union. See also https://nymag.com/intelligencer/2020/01/why-we-should-ban-facial-recognition-technology.htm; and European Commission, Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM(2021) 206 final, 21 April 2021, art. 5 (1) (d).

[50] See www.ohchr.org/Documents/Issues/Terrorism/biometricsreport.pdf.

their facial expressions, which lacks a solid scientific basis.[51] Researchers have found only a weak association of emotions with facial expressions[52] and highlighted that facial expressions vary across cultures and contexts,[53] making emotion recognition susceptible to bias and misinterpretations. Given these concerns, the use of emotion recognition systems by public authorities, for instance for singling out individuals for police stops or arrests or to assess the veracity of statements during interrogations, risks undermining human rights, such as the rights to privacy, to liberty and to a fair trial.

Artificial intelligence systems and public services

29. AI systems are increasingly being used to help deliver public services, often with the stated goal of developing more efficient systems for timely and accurate delivery of services. This is also increasingly being seen in humanitarian contexts where delivery of humanitarian goods and services may be linked to AI systems. Although these are legitimate, even laudable goals, the deployment of AI tools in the delivery of public and humanitarian services may have an adverse impact on human rights if proper safeguards are not in place.

30. AI is used in diverse public services, ranging from decision-making about welfare entitlements to flagging families for visits by childcare services.[54] These decisions are made using large data sets, which not only include State-held data but can also include information obtained from private entities, such as social media companies or data brokers, often gathered outside protective legal frameworks.[55] Furthermore, since the computational knowledge and power over AI systems tends to be held by private companies, these arrangements often mean that private companies gain access to data sets containing information about large parts of the population. This raises privacy concerns as well as concerns about how historic bias embedded in data will affect the decision-making of public authorities.

31. A major concern regarding the use of AI for public services is that it can be discriminatory, particularly with regard to marginalized groups.[56] The Special Rapporteur on extreme poverty and human rights has warned of a "digital welfare dystopia" in which unfettered data-matching is used to expose, survey and punish welfare beneficiaries and conditions are imposed on beneficiaries that undermine individual autonomy and choice.[57] These concerns were illustrated recently in the Netherlands, where a widely reported court ruling banned a digital welfare fraud detection system as it was found to infringe on the right to privacy. The system in question provided central and local authorities with broad powers to share and analyse data that were previously kept separately, including on employment, housing, education, benefits and health insurance, as well as other forms of identifiable data. Moreover, the tool targeted low-income and minority neighbourhoods, leading to de facto discrimination based on socioeconomic background.[58]

Use of artificial intelligence in the employment context

32. A range of employers across all types and sizes of business demonstrate the growing demand for monitoring and managing workers using data-driven technologies, including AI

[51] See www.nature.com/articles/d41586-020-00507-5.

[52] See https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780190613501.001.0001/acprof-9780190613501-chapter-7.

[53] See https://journals.sagepub.com/doi/10.1177/1529100619832930; and https://pubmed.ncbi.nlm.nih.gov/22509011/.

[54] See A/74/493.

[55] Submission by Privacy International.

[56] Submission by Digital Rights Watch. For an in-depth analysis of the disparate impacts of automation in welfare systems, see Virginia Eubanks, *Automating Inequality* (New York, St. Martin's Press, 2018).

[57] See A/74/493. See also the Special Rapporteur's letter IRL 1/2020, in which he noted similar concerns regarding a digital services card, and the reply thereto. Several references are made in the present report to communications sent by the special procedure mandate holders of the Human Rights Council. All such communications and the replies thereto are available from https://spcommreports.ohchr.org/Tmsearch/TMDocuments.

[58] See www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25522.

systems. So-called people analytics claims to provide more efficient and objective information about employees. This can include automated decision-making for hiring, promotion schemes or dismissal.

33. While most of the focus of such technologies lies on monitoring job-related behaviour and performance, a range of applications of AI systems also extends to non-job-related behaviour and data.[59] The COVID-19 pandemic has accelerated this trend in two ways. First, some companies that provide workers with preventive health schemes increasingly collect health-related data. Second, as more processes are executed digitally while people work from home, workplace monitoring by AI systems is taken into people's homes. Both trends increase the risk of merging the data from workplace monitoring with non-job-related data inputs. These AI-based monitoring practices constitute vast privacy risks throughout the full data life cycle. Adding to this, data can be used for other purposes than those initially communicated to employees, which can result in a so-called function creep.[60] At the same time, the quantitative social science basis of many AI systems used for people management is not solid, and is prone to biases. For example, if a company uses an AI hiring algorithm trained on historic data sets that favour male, white, middle-aged men, the resulting algorithm will disfavour women, people of colour and younger or older people who would have been equally qualified to fill the vacancy.[61] At the same time, accountability structures and transparency to protect workers are often lacking and workers are increasingly confronted with little or no explanation about AI-based monitoring practices.[62] While in some situations, companies have a genuine interest in preventing misconduct in the workplace, the measures to uphold that interest often do not justify the extensively invasive practices for quantifying the social modes of interaction and connected performance goals at work. In a workplace setting and in light of the power relationship between employer and employee, one can also envisage potential scenarios where workers are compelled to waive away their privacy rights in exchange for work.[63]

Artificial intelligence for managing information online

34. Social media platforms use AI systems to support content management decisions.[64] Companies use these systems to rank content and decide what to amplify and what to downgrade, including by personalizing these decisions to different individual users based on their profiles. Automation is also used when implementing restrictions to content, including in response to different legal requirements within and across jurisdictions.[65] The adoption of filter obligations for intermediaries relating to perceived online harms risk expanding the widespread reliance on AI without consideration of the severe impact of these systems on the rights to privacy and to freedom of expression at the local and global levels.

35. The vast data sets that curation, amplification and moderation systems rely on are created and continuously expanded through extensive online monitoring and profiling of platform users and their personal networks.[66] This perpetual process of collecting information and making inferences from it, combined with extreme market concentration, has led to the situation where a handful of companies globally hold and control profiles about billions of individuals and the networked public sphere at large.

36. AI-assisted content curation done by companies with enormous market power raises concerns about the impact on the capacity of the individual to form and develop opinions, as two successive holders of the mandate of Special Rapporteur on the promotion and protection

[59] See https://journals.sagepub.com/doi/10.1177/20539517211013051.

[60] Christl, *Corporate surveillance in everyday life*.

[61] Submission by Poland. See also www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

[62] See https://journals.sagepub.com/doi/full/10.1177/2053951720938093.

[63] See www.californialawreview.org/print/3-limitless-worker-surveillance/.

[64] See www.theverge.com/2020/11/13/21562596/facebook-ai-moderation; and https://journals.sagepub.com/doi/full/10.1177/2053951720943234.

[65] See OTH 71/2018 and OTH 73/2020. For an in-depth analysis of automated content filtering, see also https://journals.sagepub.com/doi/full/10.1177/2053951720920686.

[66] A/73/348, para. 17.

of the right to freedom of opinion and expression have pointed out.[67] Furthermore, platform recommender systems tend to focus on maximizing user engagement while relying on insights into people's preferences, demographic and behavioural patterns, which has been shown to often promote sensationalist content, potentially reinforcing trends towards polarization.[68] Moreover, the targeting of information may be unwelcome and even lead to dangerous privacy intrusions. Recommender systems have, for example, resulted in survivors of violence finding that the perpetrator was offered to them as a potential friend by social media platforms, and vice versa, putting the survivor at risk. In addition, the bias of majority or dominant groups reflected in data from search results has been shown to affect the information shared by or about minority or vulnerable groups. For example, research has demonstrated a disturbing degree of gender[69] and racial bias in Google's search results.[70]

IV. Addressing the challenges

37. The need for a human rights-based approach to new technologies in general, and artificial intelligence in particular, has been recognized by a growing number of experts, stakeholders and the international community.[71] A human rights-based approach offers a toolbox to help societies to identify ways to prevent and limit harm while maximizing the benefits of technological progress.

A. Fundamental principles

38. A human rights-based approach to AI requires the application of a number of core principles, including equality and non-discrimination, participation and accountability, principles that are also at the heart of the Sustainable Development Goals and the Guiding Principles on Business and Human Rights. In addition, the requirements of legality, legitimacy, necessity and proportionality must be consistently applied to AI technologies.[72] Moreover, AI should be deployed in a way that facilitates the realization of economic, social and cultural rights by ensuring that their key elements of availability, affordability, accessibility and quality are achieved.[73] Those who suffer human rights violations and abuses relating to the use of AI should have access to effective judicial and non-judicial remedies.[74]

39. As was pointed out above, restrictions of the right to privacy must be provided for by law, be necessary to achieve a legitimate goal, and be proportionate to that goal. In practice, that means that States are required to carefully determine if a measure is able to achieve a set objective, how important that objective is and what the impacts of the measure will be. States should also determine if less invasive approaches could achieve the same results with the same effectiveness; if so, those measures need to be taken. The High Commissioner has already outlined such necessary limitations and safeguards in the context of surveillance by intelligence agencies and law enforcement.[75] It should be noted that the necessity and proportionality tests can also lead to the conclusion that certain measures must not be taken. For example, requirements of blanket, indiscriminate retention of communications data

[67] Ibid., para. 25, and A/HRC/47/25, para. 36.

[68] See www.brookings.edu/techstream/how-youtube-helps-form-homogeneous-online-communities/.

[69] Submissions by Austria and Germany.

[70] Safiya Umoja Noble, *Algorithms of Oppression* (New York, New York University Press, 2018).

[71] General Assembly resolution 75/176, para. 6; Human Rights Council resolutions 47/16, para. 8 (d), and 47/23, sixteenth preambular para.; A/73/348, paras. 47–60, A/75/590, para. 57, and A/HRC/43/29; and submissions by Austria, Privacy Commissioner of Canada, Digital Rights Watch, Global Network Initiative and Privacy International.

[72] A/HRC/43/29, para. 41.

[73] See the detailed analysis of the role of new technologies for the realization of economic, social and cultural rights in A/HRC/43/29.

[74] International Covenant on Civil and Political Rights, art. 2 (3), and Guiding Principles on Business and Human Rights, principle 15 (c) and pillar III.

[75] A/HRC/39/29, paras. 34–41.

imposed on telecommunications and other companies would fail the proportionality test.[76] Similarly, imposing biometric identification requirements on recipients of welfare benefits is disproportionate if no alternative is provided. Moreover, it is crucial that measures are not assessed in isolation, but that the cumulative effects of distinct but interacting measures are properly taken into account. For example, before deciding to deploy new AI-based surveillance tools, a State must take stock of the existing capacities and their effects on the enjoyment of the right to privacy and other rights.

B. Legislation and regulation

40. Effective protection of the right to privacy and interlinked rights depends on the legal, regulatory and institutional frameworks established by States.[77]

41. The importance of effective legal protections under data privacy laws has grown with the emergence of data-driven AI systems. These protections should meet the minimum standards identified in the previous report of the High Commissioner on the right to privacy.[78]

42. Data privacy frameworks should account for the new threats linked to the use of AI.[79] For example, laws could impose limitations on the type of data that may legally be inferred and/or further used and shared. Legislators should also consider strengthening individuals' rights, including by granting them the rights to a meaningful explanation and to object to fully automated decisions that affect their rights.[80] As AI technologies continue to evolve, it will be necessary to continue to develop more safeguards within data privacy frameworks.

43. One key element to counter the growing complexity and opacity of the global data environment, including its vast information asymmetries, is independent data privacy oversight bodies. These bodies need to have effective enforcement powers and to be adequately resourced. Civil society organizations should be empowered to support enforcement of data privacy laws, including through the establishment of robust complaint mechanisms.

44. Beyond data privacy legislation, a broader range of laws need to be reviewed and potentially adopted to address the challenges of AI in a rights-respecting way.[81]

45. Taking into account the diversity of AI applications, systems and uses, regulation should be specific enough to address sector-specific issues and to tailor responses to the risks involved.[82] The higher the risk for human rights, the stricter the legal requirements for the use of AI technology should be. Accordingly, sectors where the stakes for individuals are particularly high, such as law enforcement, national security,[83] criminal justice, social protection, employment, health care, education and the financial sector, should have priority.

---

[76] Ibid., para. 18; and Court of Justice of the European Union, *Digital Rights Ireland and Others*, C-293/12 and C-594/12, para. 69. See also Court of Justice of the European Union, *Maximilan Schrems v. Data Protection Commissioner*, C-362/14, para. 94, finding that "legislation permitting the public authorities to have access on a generalised basis to the content of electronic communications must be regarded as compromising the essence of the fundamental right to respect for private life".

[77] A/HRC/39/29, para. 26.

[78] Ibid., paras. 28–33.

[79] The Council of Europe protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, adopted in 2018, for example, is a response to the emergence of new data-processing practices.

[80] See the General Data Protection Regulation of the European Union, which contains such rights, and the California Privacy Rights Act, which authorizes the regulator to adopt rules to that effect.

[81] See Council of Europe, Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems.

[82] The proposed AI Act of the European Union takes such a risk-based approach. The submissions from Freedom Online Coalition, Global Network Initiative and Global Partners Digital contain support for risk-based regulation.

[83] In A/HRC/27/37 and A/HRC/39/29, the High Commissioner clarified the requirements for surveillance measures taken in the context of criminal investigations and for the purposes of the protection of national security that should guide legislation in that area.

A risk-proportionate approach to legislation and regulation will require the prohibition of certain AI technologies, applications or use cases, where they would create potential or actual impacts that are not justified under international human rights law, including those that fail the necessity and proportionality tests. Moreover, uses of AI that inherently conflict with the prohibition of discrimination should not be allowed. For example, social scoring of individuals by Governments[84] or AI systems that categorize individuals into clusters on prohibited discriminatory grounds[85] should be banned in line with these principles. For systems whose use presents risks for human rights when deployed in certain contexts, States will need to regulate their use and sale to prevent and mitigate adverse human rights impacts[86] both within and outside the State's territory. Mandatory involvement of human supervision and decision-making should be prescribed when adverse human rights impacts are likely to occur.[87] Given that it can take time before risks can be assessed and addressed, States should also impose moratoriums on the use of potentially high-risk technology, such as remote real-time facial recognition, until it is ensured that their use cannot violate human rights.

46. States should also adopt robust export control regimes for the cross-border trade of surveillance technologies in order to prevent the sale of such technologies when there is a risk that they could be used for violating human rights, including by targeting human rights defenders or journalists.[88]

47. The spectrum of risks arising from AI systems suggests a need for adequate independent, impartial oversight over the development, deployment and use of AI systems. Oversight can be carried out by a combination of administrative, judicial, quasi-judicial and/or parliamentary oversight bodies.[89] For example, in addition to data privacy authorities, consumer protection agencies, sectoral regulators, anti-discrimination bodies and national human rights institutions should form part of the oversight system. Moreover, cross-sectoral regulators dedicated to overseeing the use of AI can help to set fundamental standards and ensure policy and enforcement coherence.

C. Human rights due diligence

48. States and businesses should ensure that comprehensive human rights due diligence is conducted when AI systems are acquired, developed, deployed and operated, as well as before big data held about individuals are shared or used. [90] As well as resourcing and leading such processes, States may also require or otherwise incentivize companies to conduct comprehensive human rights due diligence.

---

[84] Submission by European Union; Catelijne Muller, "The impact of artificial intelligence on human rights, democracy and the rule of law", report to the Council of Europe, Ad hoc Committee on Artificial Intelligence (CAHAI(2020)06-fin), 24 June 2020, para. 75; and United Nations Educational, Scientific and Cultural Organization (UNESCO), "Draft text of the recommendation on the ethics of artificial intelligence" (SHS/IGM-AIETHICS/2021/JUN/3 Rev.2), 25 June 2021, para. 26.

[85] See European Data Protection Board and European Data Protection Supervisor, joint opinion 5/2021, para. 33.

[86] Submission by Derechos Digitales.

[87] See www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6.

[88] A/HRC/41/35, para. 49, and A/HRC/44/24 para. 40. In those reports, the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression and the High Commissioner also called for a moratorium on granting export licences for surveillance technologies.

[89] See https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

[90] The B-Tech project of the Office of the United Nations High Commissioner for Human Rights (OHCHR) is developing guidance on the implementation of the Guiding Principles on Business and Human Rights in the technology industry, including responses to the human rights impact of the use of AI technologies. See www.ohchr.org/EN/Issues/Business/Pages/B-TechProject.aspx.

49. The aim of human rights due diligence processes is to identify, assess, prevent and mitigate adverse impacts on human rights that an entity may cause or to which it may contribute or be directly linked.[91] Where due diligence processes reveal that a use of AI is incompatible with human rights, due to a lack of meaningful avenues to mitigate harms, this form of use should not be pursued further. Assessing human rights impacts is an essential element of human rights due diligence processes.[92] Due diligence should be conducted throughout the entire life cycle of an AI system.[93] Particular attention should be paid to disproportionate impacts on women and girls, lesbian, gay, bisexual, transgender and queer individuals, persons with disabilities, persons belonging to minorities, older persons, persons in poverty and other persons who are in a vulnerable situation.

50. Meaningful consultations should be carried out with potentially affected rights holders and civil society, while experts with interdisciplinary skills should be involved in impact assessments, including in the development and evaluation of mitigations. States and businesses should continuously monitor the impacts of the AI systems they use to verify whether they have adverse human rights impacts. The results of human rights impact assessments, action taken to address human rights risks and public consultations should themselves be made public.[94]

D. State-business nexus

51. Situations where there is a close nexus between a State and a technology company require dedicated attention.[95] The State is an important economic actor that can shape how AI is developed and used, beyond the States' role in legal and policy measures. Where States work with AI developers and service providers from the private sector, States should take additional steps to ensure that AI is not used towards ends that are incompatible with human rights. Such steps should be applied across the management of State-owned companies, research and development funding, financial and other support provided by States to AI technology companies, privatization efforts and public procurement practices.

52. Where States operate as economic actors, they remain the primary duty bearer under international human rights law and must proactively meet their obligations. At the same time, businesses remain responsible for respecting human rights when collaborating with States and should seek ways to honour human rights when faced with State requirements that conflict with human rights law.[96] For example, when faced with demands for access to personal data that fail to meet human rights standards, they should use their leverage to resist or mitigate the harm that could be caused.[97]

53. States can enhance human rights protections by consistently requiring responsible business conduct. For example, when export credit agencies offer support to AI technology companies, they should ensure that these companies have a robust track record in rights-respecting conduct and can demonstrate this through robust due diligence processes.

54. When States rely on AI businesses to deliver public goods or services, States must ensure that they can oversee the development and deployment of the AI systems. This can be done by demanding and assessing information about the accuracy and risks of an AI application. Where risks cannot be effectively mitigated, States should not use AI to deliver public goods or services.

---

[91] For an overview of human rights due diligence in the context of AI, see https://international-review.icrc.org/articles/ai-humanitarian-action-human-rights-ethics-913, pp. 174–178.

[92] For a concise summary of human rights impact assessment methodologies, see https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

[93] A/HRC/43/29, para. 62 (g), and A/HRC/44/24, paras. 38, 53 (j) (i) and 54 (c).

[94] A/73/348, para. 68.

[95] See www.ohchr.org/Documents/Issues/Business/B-Tech/b-tech-foundational-paper-state-duty-to-protect.pdf.

[96] Guiding Principles on Business and Human Rights, principle 23 (b).

[97] A/HRC/32/38, para. 58. See also www.ohchr.org/Documents/Issues/Terrorism/biometricsreport.pdf, pp. 39–40.

E. Transparency

55. Developers, marketers, operators and users of AI systems should drastically increase their efforts regarding transparency around the use of AI. As a first step, States, businesses and other users of AI should make information available about the kind of systems they use, for what purposes, and the identity of the developer and operator of the systems.[98] Affected individuals should systematically be informed when decisions are being or have been made automatically or with the help of automation tools.[99] Individuals should also be notified when the personal data they provide will become part of a data set used by an AI system.[100] Moreover, for human rights-critical applications, States should introduce registers containing key information about AI tools and their use.[101] Effective enforcement of transparency obligations and data access, erasure and rectification rights contained in data privacy frameworks should be ensured. Particular attention should be given to enabling individuals to better understand and control the profiles compiled about them.[102]

56. Promoting transparency should go further by including sustained efforts to overcome the "black box" problem described above. The development and systematic deployment of methodologies to make AI systems more explainable – often referred to as algorithmic transparency – is of utmost importance for ensuring adequate rights protections.[103] This is most essential when AI is used to determine critical issues within judicial processes or relating to social services that are essential for the realization of economic, social and cultural rights. Researchers have already developed a range of approaches that further that goal,[104] and increased investments in this area are essential. States should also take steps to ensure that intellectual property protections do not prevent meaningful scrutiny of AI systems that have human rights impacts.[105] Procurement rules should be updated to reflect the need for transparency, including auditability of AI systems.[106] In particular, States should avoid using AI systems that can have material adverse human rights impacts but cannot be subject to meaningful auditing.[107]

V. Conclusions and recommendations

A. Conclusions

57. **The present report has highlighted the undeniable and steadily growing impacts of AI technologies on the exercise of the right to privacy and other human rights, both for better and for worse. It has pointed to worrying developments, including a sprawling ecosystem of largely non-transparent personal data collection and exchanges that underlies parts of the AI systems that are widely used. These systems affect**

98 A/HRC/43/29, para. 52, and A/73/348, para. 49.

99 Council of Europe, "Guidelines on addressing the human rights impacts of algorithmic systems", (appendix to Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems), sect. B, para. 4.2.

100 A/73/348, para. 49.

101 A/HRC/43/29, para. 52. The European Union proposal for an AI Act contains provisions on a register for high-risk AI systems.

102 See https://link.springer.com/article/10.1007/s12394-008-0003-1, p. 67.

103 For an overview of elements of algorithmic transparency, see www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6, pp. 320–323. See https://arxiv.org/abs/2001.00973 and https://arxiv.org/pdf/1711.01134.pdf.

105 Council of Europe, "Guidelines on addressing the human rights impacts of algorithmic systems", (appendix to Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems), sect. B, para. 4.1.

106 See submissions by Germany, Derechos Digitales, Freedom Online Coalition and Global Partners Digital.

107 A/73/348, para. 55, and A/HRC/43/29, para. 54.

government approaches to policing and the administration of justice, determine the accessibility of public services, decide who has a chance to be recruited for a job, and affect what information people see and can share online. Moreover, the risk of discrimination linked to AI-based decisions is all too real. The report outlines a range of ways to address the fundamental problems associated with AI, underscoring that only a comprehensive human rights-based approach can ensure sustainable solutions to the benefit of all.

58. **Nevertheless, given the diversity of new questions arising in the context of AI, the present report is a snapshot of the constantly evolving AI landscape. Areas that deserve further analysis include health, education, housing and financial services. Biometric technologies, which are becoming increasingly a go-to solution for States, international organizations and technology companies, are an area where more human rights guidance is urgently needed. Furthermore, one focus of future work from a human rights perspective should be on finding ways to fill the immense accountability gap in the global data environment. Lastly, solutions for overcoming AI-enabled discrimination should urgently be identified and implemented.**

B. Recommendations

59. **The High Commissioner recommends that States:**

(a) **Fully recognize the need to protect and reinforce all human rights in the development, use and governance of AI as a central objective, and ensure equal respect for and enforcement of all human rights online and offline;**

(b) **Ensure that the use of AI is in compliance with all human rights and that any interference with the right to privacy and other human rights through the use of AI is provided for by law, pursues a legitimate aim, complies with the principles of necessity and proportionality and does not impair the essence of the rights in question;**

(c) **Expressly ban AI applications that cannot be operated in compliance with international human rights law and impose moratoriums on the sale and use of AI systems that carry a high risk for the enjoyment of human rights, unless and until adequate safeguards to protect human rights are in place;**

(d) **Impose a moratorium on the use of remote biometric recognition technologies in public spaces, at least until the authorities responsible can demonstrate compliance with privacy and data protection standards and the absence of significant accuracy issues and discriminatory impacts, and until all the recommendations set out in A/HRC/44/24, paragraph 53 (j) (i–v), are implemented;**

(e) **Adopt and effectively enforce, through independent, impartial authorities, data privacy legislation for the public and private sectors as an essential prerequisite for the protection of the right to privacy in the context of AI;**

(f) **Adopt legislative and regulatory frameworks that adequately prevent and mitigate the multifaceted adverse human rights impacts linked to the use of AI by the public and private sectors;**

(g) **Ensure that victims of human rights violations and abuses linked to the use of AI systems have access to effective remedies;**

(h) **Require adequate explainability of all AI-supported decisions that can significantly affect human rights, particularly in the public sector;**

(i) **Enhance efforts to combat discrimination linked to the use of AI systems by States and business enterprises, including by conducting, requiring and supporting systematic assessments and monitoring of the outputs of AI systems and the impacts of their deployment;**

(j) **Ensure that public-private partnerships in the provision and use of AI technologies are transparent and subject to independent human rights oversight, and do not result in abdication of government accountability for human rights.**

60. **The High Commissioner recommends that States and business enterprises:**

(a) **Systematically conduct human rights due diligence throughout the life cycle of the AI systems they design, develop, deploy, sell, obtain or operate. A key element of their human rights due diligence should be regular, comprehensive human rights impact assessments;**

(b) **Dramatically increase the transparency of their use of AI, including by adequately informing the public and affected individuals and enabling independent and external auditing of automated systems. The more likely and serious the potential or actual human rights impacts linked to the use of AI are, the more transparency is needed;**

(c) **Ensure participation of all relevant stakeholders in decisions on the development, deployment and use of AI, in particular affected individuals and groups;**

(d) **Advance the explainability of AI-based decisions, including by funding and conducting research towards that goal.**

61. **The High Commissioner recommends that business enterprises:**

(a) **Make all efforts to meet their responsibility to respect all human rights, including through the full operationalization of the Guiding Principles on Business and Human Rights;**

(b) **Enhance their efforts to combat discrimination linked to their development, sale or operation of AI systems, including by conducting systematic assessments and monitoring of the outputs of AI systems and of the impacts of their deployment;**

(c) **Take decisive steps in order to ensure the diversity of the workforce responsible for the development of AI;**

(d) **Provide for or cooperate in remediation through legitimate processes where they have caused or contributed to adverse human rights impacts, including through effective operational-level grievance mechanisms.**