Privacy International's submission for the UN High Commissioner for Human Rights' report on the right to privacy and artificial intelligence

May 2021

1. Introduction

Privacy International (PI)[1] welcomes the opportunity to provide input to the forthcoming report by the High Commissioner for Human Rights (HCHR) on the right to privacy and artificial intelligence (AI.)[2]

AI and its applications are becoming a part of everyday life: social media newsfeeds, mediating traffic flow in cities, connected consumer devices, automated cars, eligibility mechanisms for welfare services, access to medical diagnostics, location tracking, spam filters, voice recognition systems, and search engines. If implemented responsibly, AI has the potential to promote the enjoyment of human rights. However, there is a real risk that commercial and state use has a detrimental impact on human rights.

PI believes that the HCHR thematic report is an important opportunity to build upon the analysis developed at the expert seminar in May 2020 in order to clarify states' obligations and companies' responsibilities under international human rights law in relation to the planning, design, use and assessment of AI applications.

In particular PI suggests the following main aspects should be covered in the HCHR report:

- Reassert that any interference with the right to privacy due to the use of AI technologies should be subject to the overarching principles of legality, necessity and proportionality.
- Establish the need for a human rights-based approach to all AI applications and describe the necessary measures to achieve it (including human rights by design and human rights impact assessments).
- Identify the human rights risks of specific AI applications, due to the technologies employed and/or the context of their use; and describe the circumstances when AI applications should be banned because of human rights concerns.

[1] PI is an international non-governmental organisation, which campaigns against companies and governments who exploit individuals' data and technologies. PI employs specialists in their fields, including technologists and lawyers, to understand the impact of existing and emerging technology upon data exploitation and our right to privacy.
[2] OHCHR, Call for input: report on "the right to privacy in the digital age", https://www.ohchr.org/EN/Issues/DigitalAge/Pages/cfi-digital-age.aspx.

- Encourage states to adopt or review effective data protection legislation and sectoral laws to address the negative human rights implications of AI applications – at individual, group and society level.
- Note that states have a responsibility to respect and protect human rights from threats arising by the use of AI technologies. On the one hand, state regulation can shape how the private sector develops and applies AI systems and technologies. On the other hand, states have a responsibility to ensure that public sector uses of AI – particularly in health care, welfare, migration, policing, and surveillance, is used responsibly.
- Define the scope of responsibility of non state actors, including companies and international organisations, for AI uses and the need for mechanisms to ensure that they are held accountable.

2. Key concerns regarding AI and the right to privacy

PI has long documented how AI applications are often used to process data and to identify individuals, predict and influence their behaviours. In particular AI technologies are being used:

- to infer and generate sensitive information about people;
- to profile people based upon population-scale data;
- to identify people who wish to remain anonymous; and
- to make decisions on the basis of the analysis of this data.

AI-driven consumer products and autonomous systems are frequently equipped with sensors that generate and collect vast amounts of data without the knowledge or consent of the users or those in their proximity.[3] On the internet, vast amounts of data about people's lives and behaviour is increasingly gathered through tracking technologies, including sensitive data, for example on mental health websites[4] or menstruation apps.[5] AI applications facilitate the further analysis of this data and the generation of inferences to create finely grained profiles.[6] Such profiles are then used to target people with advertising – both commercial and political – and ultimately feed into other consequential decisions which may negatively affect human rights, including access to credit and insurance. AI applications are also increasingly being used in digital identity systems for a range of purposes, including authentication and verification.[7]

[3] For further information on each of these, see PI and ARTICLE 19 publication on "Privacy and Freedom of Expression In the Age of Artificial Intelligence", April 2018, available at https://privacyinternational.org/sites/default/files/2018-04/Privacy%20and%20Freedom%20of%20Expression%20%20In%20the%20Age%20of%20Artificial%20Intelligence.pdf , pp. 6-7.
[4] PI, "Your mental health for sale", https://privacyinternational.org/campaigns/your-mental-health-sale.
[5] PI, "No Body's Business But Mine: How Menstruation Apps Are Sharing Your Data", https://privacyinternational.org/long-read/3196/no-bodys-business-mine-how-menstruations-apps-are-sharing-your-data.
[6] There is an entire ecosystem dedicated to these privacy invasive practices, including data brokers and ad tech companies. PI, "Challenge to Hidden Data Ecosystem", https://privacyinternational.org/legal-action/challenge-hidden-data-ecosystem
[7] For example, Yoti, developed Yoti Age Scan technology, that uses AI to estimate an individual's age based on their image. This is used, for example, within the Yoti app instead of providing a verified ID document that contains their age in order to be able to buy alcohol or to access adult content online. For further information see: PI, "The Identity Gatekeepers and the Future of Digital Identity", https://privacyinternational.org/long-read/3254/identity-gatekeepers-and-future-digital-identity and PI "Demanding identity systems on our terms", https://staging.privacyinternational.org/campaigns/demanding-identity-systems-our-terms

As such, AI applications can affect the whole range of human rights.[8] Because of the central role data play in most AI applications, the right privacy is particularly affected. Some of the key concerns regarding AI applications and privacy are:

- **Data exploitation:** AI applications frequently rely on the generation, collection, processing, and sharing of large amounts of data, both about individual and collective behaviour. This data can be used to profile individuals and predict future behaviour. It is often difficult to fully understand what kinds and how much data devices, networks, and platforms generate, process, or share, indeed this is often opaque by design.
- **Opacity and secrecy of profiling and automated decision making:** Some AI applications can be opaque to individuals, regulators, or even the designers of the system themselves, making it difficult to challenge or interrogate outcomes. While there are technical solutions to improving the interpretability and/or the ability to audit of some systems for different stakeholders, a key challenge remains where this is not possible, and the outcome has significant impacts on people's lives.
- **Re-identification and de-anonymisation:** AI applications can be used to identify and thereby track individuals across different devices, in their homes, at work, and in public spaces. For example, while personal data is routinely (pseudo-) anonymised within datasets, AI can be employed to de-anonymise this data.[9]
- **Discrimination, unfairness, inaccuracies and bias:** AI-driven identification, profiling, and automated decision-making may also lead to unfair, discriminatory, or biased outcomes. People can be misclassified, misidentified, or judged negatively, and such errors or biases may disproportionately affect certain groups of people.

3. AI applications and contexts of particular concerns

The term 'Artificial Intelligence' or 'AI' is used to refer to a diverse range of applications and technologies, with different levels of complexity, autonomy and abstraction. This broad usage encompasses machine learning (which makes inferences, predictions and decisions about individuals), domain-specific AI algorithms, fully autonomous and connected objects and even the futuristic idea of an AI 'singularity'. This lack of definitional clarity is a challenge: different types of AI applications and the context into which they are deployed raise specific regulatory issues.[10]

Without aiming to be comprehensive, in the following sections PI describes how specific AI applications and AI applications in specific sectors negatively affect the enjoyment of the right to privacy and other human rights.

---

[8] As noted by the UN General Assembly resolution on the right to privacy in the digital age, "artificial intelligence or machine-learning technologies [...] may lead to decisions that have the potential to affect the enjoyment of human rights, including economic, social and cultural rights, and affect non-discrimination".(The right to privacy in the digital age, GA Res 75/176, 16 December 2020, https://undocs.org/A/RES/75/176.) See also PI, "Artificial Intelligence", Explainer, https://privacyinternational.org/learn/artificial-intelligence.

[9] Multiple studies have shown that potential de-anonymisation capabilities of AI technologies. Similarly, in a more recent study published in Nature, researchers were able to demonstrate that, despite the anonymisation techniques applied, "data can often be reverse engineered using machine learning to re-identify individuals." Luc Rocher, Julien M. Hendrickx & Yves-Alexandre de Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative models", 23 July 2019, https://www.nature.com/articles/s41467-019-10933-3.

[10] On definitions of different AI applications and techniques, see PI and ARTICLE 19 publication on "Privacy and Freedom of Expression In the Age of Artificial Intelligence", April 2018, available at https://privacyinternational.org/sites/default/files/2018-04/Privacy%20and%20Freedom%20of%20Expression%20%20In%20the%20Age%20of%20Artificial%20Intelligence.pdf, pp. 6-7.

3.1 AI and facial recognition technology

Facial recognition technology (FRT) typically refers to systems which collect and process data about a person's face. FRT can be used for the purposes of identification, authentication/verification or categorisation of those individuals. Such technologies are highly intrusive because they rely on the capture, extraction, storage or sharing of people's biometric facial data.[11]

As FRT has the power to fundamentally change the very meaning of public space and anonymity both online and off-line, PI believes that the deployment of this technology should be approached with great caution and it should be seriously considered whether the use of FRT is permissible at all in light of the obligations imposed by international human rights law.

The HCHR and UN human rights experts and bodies have expressed significant concerns about the use of FRT particularly its use to monitor assemblies.[12]

PI is also concerned that FRT for identification and categorisation purposes could lead to discrimination. FRT relies on probabilistic reasoning, and as such, inevitably produces varying levels of false positive and false negatives. Many commercially available facial recognition systems have been found to have different error rates, depending on people's race and gender.[13] In his 2019 Report, the UN Special Rapporteur on the right to freedom expression noted that FRT "seeks to capture and detect the facial characteristics of a person, potentially profiling individuals based on their ethnicity, race, national origin, gender and other characteristics, which are often the basis for unlawful discrimination".[14]

Based on our research and analysis, PI believes that live FRT in public places by state and non-state actors should be banned. The introduction of live FRT would result in the normalisation of surveillance across all societal levels and accordingly cast a "chilling effect" on the exercise of fundamental rights, such as our freedom of expression and freedom of assembly. Live FRT casts a chilling effect on societies and impose a sense of constant surveillance, self-restriction and self-censoring, and normalises indiscriminate surveillance.

PI recognises that in limited circumstances and subject to strict safeguards, the deployment of static FRT by state actors such as law enforcement agencies could be justified. PI has highlighted the specific conditions in accordance with international human rights law on which any decision to use FRT technology should depend in its submission to the Scottish Parliament.[15] In summary, the minimum safeguards should include strict

[11] FRT may involve the use of cameras, which can capture individuals' facial images and process them in real time ("live FRT") or at a later point ("Static" or "Retrospective FRT"). The collection of facial images results in the creation of "digital signatures of identified faces", which are analysed against one or more databases ("Watchlists"), usually containing facial images obtained from other sources to determine if there is a match.
[12] See, for example, High Commissioner for Human Rights, report on the Impact of new technologies on the promotion and protection of human rights in the context of assemblies, including peaceful protests, UN doc. A/HRC/44/24. The UN Special Rapporteur on freedom of opinion and expression has called for a moratorium of the sale and use of live facial recognition (LFR) technology (Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, A/HRC/41/35, 28 May 2019, para 66ff.)
[13] See Karen Hao, AI is sending people to jail—and getting it wrong (MIT Technology Review, 21 January 2019) https://www.technologyreview.com/s/612775/algorithms-criminal-justice-ai/
[14] Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (A/HRC/41/35, 28 May 2019), para 12.
[15] On a complete analysis of facial recognition concerns, see PI, "Submission to the Scottish Parliament's Justice Sub-Committee on Policing's inquiry into facial recognition policing", November 2019,

application of the principles of legality, necessity and proportionality, prior judicial authorisation on the basis of reasonable suspicion of serious crime or serious threat to national security, strict rules on retention and destruction of personal data, prior judicial authorisation and independent monitoring and oversight, transparency in relation to the criteria used for the inclusion of individuals into watchlists, access to effective remedies, including the rights of individuals to be adequately notified of the processing of their biometric data and be given the opportunity to exercise their rights of rectification, access, erasure, as well as to challenge any processing operation before competent courts and regulators.

The use of FRT by private companies, such as the processing of facial images of people entering retailers and hospitality venues, or spaces owned by private actors, is seriously questionable whether the use of such an intrusive technology can be considered compliant with the principles enshrined in data protection and privacy law. As for the processing of facial images obtained online, please see the following section.

3.2 AI and Social Media Intelligence (SOCMINT)

Over the last few years, governments and companies have significantly developed their capacity to carry out social media intelligence (SOCMINT), the techniques and technologies that allow them to monitor social media networking sites, such as Facebook, Instagram, YouTube or Twitter.[16]

These activities are undertaken directly by government themselves but in some instances, governments are calling on companies to provide them with the tools and/or knowhow to undertake these activities.[17] For example, companies like Clearview AI trawl through sites like Instagram, YouTube and Facebook, as well as personal blogs and professional websites, and save a copy of public photos that contain a face. Clearview AI claims to have "the largest known database of 3+ billion facial images". They then use FRT (see above) to extract the unique features of people's faces, effectively building a gigantic database of our biometrics.[18]

The collection and analysis of *publicly available* content on social media without informed public awareness and debate, clear and precise legal frameworks, and robust safeguards fall short of standards of protection of the right to privacy and of personal data protection. Governments and companies often argue that this collection and analysis of data obtained from social media have little impact on people's privacy as and when it relies "only" on *publicly available* information. This inaccurate representation fails to account for

---

https://privacyinternational.org/advocacy/3274/submission-scottish-parliaments-justice-sub-committee-policing-inquiry-facial

[16] See PI, "Social Media Intelligence", 23 October 2017, https://privacyinternational.org/explainer/55/social-media-intelligence

[17] For example, the company, Giant Oak, markets itself to government and financial institutions and describes its Giant Oak Search Technology (GOST) as an "open source search and triage tool" that leverages open sources, social media, and the deep web to identify evidence of illicit activity and relevant information about entities of interest to clients. Their tool scrapes the web to pull in and search through vast amounts of information available online – such as news stories, blog posts, and images – as well as social media information. Layered on top of the search capability, the tool uses "sophisticated analytics scoring" to prioritise how results are shown, allows customers to search by key words, and provides a "dossier creation user interface" See PI, "Who Supplies the Data, Analysis, and Tech Infrastructure to US Immigration Authorities?", 9 August 2018, https://privacyinternational.org/long-read/2216/who-supplies-data-analysis-and-tech-infrastructure-us-immigration-authorities

[18] In May 2021 PI and three other organisations filed a series of legal complaints against Clearview AI, Inc. to data protection regulators in France, Austria, Italy, Greece and the United Kingdom. See PI: https://privacyinternational.org/legal-action/challenge-against-clearview-ai-europe

the intrusive nature of collection, retention, use, and sharing of a person's personal data obtained from public places and through social media. By way of example, 'tweets' posted from a mobile phone can reveal location data, and their content can also reveal individual opinions (including political opinions) as well as information about a person's preferences, sexuality, and health status. This privacy invasion is made possible by the development of AI technologies that automatically process and aggregate a vast range of data.

PI is concerned that the practice of social media monitoring described above is mostly carried out without appropriate legal frameworks or remedies. The data collected feeds AI applications that are used in a variety of context, from predictive policing[19], to monitoring migration routes[20], to investigating fraud of welfare services and other minor offences.[21] Such AI applications can amplify discriminatory and abusive practices against specific groups in the population.[22] For example, following freedom of information requests, PI found that in the UK, local government authorities are looking at people's social media accounts, such as Facebook, as part of their intelligence gathering and investigation tactics in areas such as council tax payments, children's services, benefits and monitoring protests and demonstrations. In some cases, local authorities will go so far as to use such information to make accusations of fraud and withhold urgently needed support from families who are living in extreme poverty.[23]

3.4 Applications of AI technology negatively affecting the most vulnerable groups

In recent years, PI has exposed the negative effect of AI applications on some of the most vulnerable groups in society, documenting how the use of AI has exacerbated, rather than addressed, existing discrimination and exclusion.[24]

- AI in welfare

Current and emerging AI supported processes to access social welfare are designed and managed in a way that comes at the cost of everyone's privacy, dignity and autonomy. From the stage of eligibility and registration to access benefits, recipients need to turn over vast amounts of personal data – about their employment, their health conditions, their relationship status – which is processed by AI applications to make (or support the making of) decision related to access to social welfare benefits.[25] Governments across the world are building technologically integrated programmes to allow individuals to access welfare payments. Social protection systems around the world are increasingly 'conditional', meaning that aspects of state support, usually financial or practical, are dependent on

[19] PI, "How predictive policing technology can lead to discrimination and profiling", https://privacyinternational.org/node/2720.
[20] PI, "Who supplies the data, analysis, and tech infrastructure to US immigration authorities?", 9 August 2018, https://privacyinternational.org/long-read/2216/who-supplies-data-analysis-and-tech-infrastructure-us-immigration-authorities
[21] PI, "Shedding light on the DWP Part 2 - A Long Day's Journey Towards Transparency", 14th February 2021, https://privacyinternational.org/long-read/4397/shedding-light-dwp-part-2-long-days-journey-towards-transparency
[22] The Fundamental Rights Agency (FRA) further points out that the number of stops and checks by the police might increase, because of the larger number of hits they get from true and false positives. Therefore, the risk of inappropriate police behaviour would increase due to additional stress. European Union Agency for Fundamental Rights, facial recognition technology: fundamental right considerations in the context of law enforcement, https://fra.europa.eu/en/publication/2019/facial-recognition-technology-fundamental-rights-considerations-context-law, p. 20.
[23] See https://privacyinternational.org/report/3584/when-local-authorities-arent-your-friends
[24] Examples of abuse of AI applications can be found here https://privacyinternational.org/examples?field_type_of_abuse_target_id_2%5B%5D=264.
[25] PI, "When Big Brother Pays Your Benefits", https://privacyinternational.org/taxonomy/term/675.

claimants complying with a set of rules or conditions. These processes are increasingly tied to rigid digital identification systems and determined by algorithmic and automated decision making processes.[26] Those who fail to comply with the rules can find themselves automatically cut-off, have their assistance reduced or are subject to sanctions and fines. In some cases the most vulnerable groups of the population are subject to particularly intrusive level of control and surveillance via digital technologies.[27] PI and its partner organisations have seen the involvement of industry in enabling such practices[28] with companies like IBM offering a wider range of 'solutions' for automating social benefits systems including child welfare programmes[29], and data brokers like Experian being used in welfare programmes like Colombia' System of Identification of Social Program Beneficiaries (SISBÉN)[30]. Concerns about the negative impact of the use of AI applications in the welfare context have already been expressed by UN human rights experts[31] and national courts are beginning to rule against these systems on the grounds that they fail to comply with human rights law.[32]

- **AI in immigration enforcement and border control**

New technologies have been deployed in immigration enforcement including AI and automated decision making.[33] These have included lie detectors at the border,[34] tracking of social media accounts,[35] language analysis[36], automated decision making of about visitor visa applications[37], to the identification refugees,[38] or as part of digital border monitoring systems.[39] There is often no or inadequate legal framework regulating the deployment of these technologies by public authorities and private security companies and in most cases there are not effective safeguards to protect refugee and migrants

[26] See https://privacyinternational.org/news-analysis/3112/stage-1-applying-social-benefits-facing-exclusion
[27] See https://privacyinternational.org/explainer/4425/what-aspen-card-and-why-does-it-need-reform
[28] See https://privacyinternational.org/long-read/4144/benefitting-whom-overview-companies-profiting-digital-welfare
[29] See https://www.ibm.com/products/watson-health-child-welfare
[30] See https://web.karisma.org.co/experimentar-con-los-datos-de-personas-en-situacion-de-pobreza-una-mala-practica-para-lograr-la-justicia-social-en-colombia/
[31] See Report of the Special rapporteur on extreme poverty and human rights", UN doc. A/74/48037, 11 October 2019.
[32] In a landmark ruling, a Dutch has now concluded that the use of SyRI is unlawful as it violates the right to privacy. The court found that the Dutch government had failed to strike a balance between the right to privacy and the public interest in detecting welfare fraud, and that the use of SyRI was disproportionate to the aim it sought to achieve. See PI, "The SyRI case: a landmark ruling for benefits claimants around the world", 5 February 2020, https://privacyinternational.org/news-analysis/3363/syri-case-landmark-ruling-benefits-claimants-around-world.
[33] See PI's work on demanding a human approach to immigration, available at: https://privacyinternational.org/what-we-do/demand-humane-approach-immigration; "PI's submission to the 'UN Working Group on the use of mercenaries' on the role of private companies in immigration and border management and the impact on the rights of migrants", https://privacyinternational.org/advocacy/3756/pis-submission-un-working-group-use-mercenaries-role-private-companies-immigration, 07 May 2020.
[34] iborderCtrl website, https://www.iborderctrl.eu/The-project.
[35] PI, '#PrivacyWins: EU Border Guards Cancel Plans to Spy on Social Media (for now)', https://privacyinternational.org/advocacy/3289/privacywins-eu-border-guards-cancel-plans-spy-social-media-now
[36] PI, 'The UK's Privatised Migration Surveillance Regime: A rough guide for civil society', 2021, https://privacyinternational.org/sites/default/files/2021-01/PI-UK_Migration_Surveillance_Regime.pdf
[37] Foxglove, "Legal action to challenge Home Office use of secret algorithm to assess visa applications", https://www.foxglove.org.uk/news/legal-challenge-home-office-secret-algorithm-visas.
[38] Patrick Tucker, "Refugee or Terrorist? IBM thinks its software has the answer", *Defense One*, 27 January 2016, http://www.defenseone.com/technology/2016/01/refugee-or-terrorist-ibm-thinks-its-software-has-answer/125484/.
[39] Olivia Solon, "'Surveillance society': has technology at the US-Mexico border gone too far?", *The Guardian*, 13 July 2018, https://www.theguardian.com/technology/2018/jun/13/mexico-us-border-wall-surveillance-artificial-intelligence-technology.

against undue interferences with their privacy. Because of their heightened vulnerability, refugee and migrants are very unlikely to be in a position to object to the application of these technologies or to seek remedy against abuses.

3.5 AI in Covid-19 pandemic responses

Since the start of the pandemic the use of AI has been observed for a variety of purposes and in various sectors from medical to law enforcement, including to:

- study the virus and research for a vaccine, early medical diagnosis[40], prediction, and modelling of spread and future outbreaks, such early warning alerts,[41] analyse public health impacts[42];
- predict and track people who have contracted the virus and who might develop respiratory problems[43];
- contact tracing[44], and social control[45] as seen with the use of AI to oversee enforcement of quarantine and other measures of social control[46].

Despite the hype around the beneficial use of AI to fight the pandemic, doubts emerged early on particularly as to how much AI can support public health efforts.[47] Whilst opportunities have emerged particularly for the health and epistemological community these are still nascent, i.e. pilot and not scalable yet, and there is still limited evidence of the results,[48] and there is growing awareness of the conditions needed for AI to be beneficial and effective, including "the need for good quality and flow of data", while being cautious about "the ethical concerns, i.e. trust and privacy", that are trigged with the use of AI.[49]

The medical research community has also challenged the lack of transparency and regulatory void in which these AI technologies are being deployed, and how this could lead to more harm than good. Medical researchers have also called for the design of AI models to be done in collaboration with healthcare workers to understand how these could be applied in practice and with what implications.[50]

---

[40] Jane Wakefield, "Coronavirus: AI steps up in battle against Covid-19", *BBC News*, 18 April 2020, https://www.bbc.co.uk/news/technology-52120747.
[41] John McCormick, "Online Map Tracks Coronavirus Outbreak in Real Time", *The Wall Street Journal,* 5 March 2020, https://www.wsj.com/articles/online-map-tracks-coronavirus-outbreak-in-real-time-11583354911
[42] See: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7692869/; https://link.springer.com/article/10.1007/s10796-021-10131-x; https://www.coe.int/en/web/portal/covid-19-artificial-intelligence; https://ec.europa.eu/digital-single-market/en/content/digital-technologies-actions-response-coronavirus-pandemic-data-artificial-intelligence-and
[43] "How Artificial Intelligence is helping the fight against COVID-19", *Health Europa*, 8 April 2020, https://www.healtheuropa.eu/how-artificial-intelligence-is-helping-the-fight-against-covid-19/99258/.
[44] John McCormick, "Online Map Tracks Coronavirus Outbreak in Real Time", *The Wall Street Journal,* 5 March 2020, https://www.wsj.com/articles/online-map-tracks-coronavirus-outbreak-in-real-time-11583354911
[45] See Wim Naudé, Artificial Intelligence against Covid-19 – an early review, April 2020, http://ftp.iza.org/dp13110.pdf
[46] "Russian centre uses AI and cameras to curb misinformation and monitor quarantines", https://privacyinternational.org/examples/3485/russian-centre-uses-ai-and-cameras-curb-misinformation-and-monitor-quarantines; See "Social control" http://ftp.iza.org/dp13110.pdf
[47] See: https://www.wired.com/story/artificial-intelligence-couldnt-save-us-from-covid-19/
[48] See Wim Naudé, Artificial Intelligence against Covid-19 – an early review, April 2020, http://ftp.iza.org/dp13110.pdf
[49] "Discussion: Where we are and What is Next" and "Ethical Aspects" https://link.springer.com/article/10.1007/s10796-021-10131-x
[50] The lancet, Artificial intelligence for COVID-19: saviour or saboteur?, January 2021, https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30295-8/fulltext

Significant human rights concerns have emerged in the context of using AI for social control measures, such as quarantine enforcement.[51] Such measures have ranged from the use of thermal images and drones[52] to location tracking and facial recognition. Information gathering and analysis, including by companies like Palantir[53] providing AI data-driven solutions to governments, including the UK, to process health data of millions of users without any guarantees on who has access to this data or what future applications they may be used for.

PI is concerned that measures, including AI applications, adopted in the context of the public health emergency caused by the COVID-19 pandemic may remain in place indefinitely and become the foundation for policing and law enforcement strategies.[54]

These ongoing concerns highlight the need for careful consideration before deploying AI technologies, for clearer and enforceable regulatory mechanisms[55] as well as for wider consultation between designers, often in the private sector, and the health and medical community where these tools are intended to be deployed to understand the risks and benefits.[56]

4. Assessing the national legal frameworks

The overarching principles of legality, necessity and proportionality apply to any use of AI technology that interferes with the right to privacy. The data protection legal framework – requiring *inter alia* an appropriate legal basis for any data processing, fairness and transparency, ensuring purpose limitation and data minimisation, accuracy, storage limitation, integrity and security, and accountability[57] - should apply to any application of AI technology that process personal data, whether used by governments or private actors.

In practice, however, the data protection frameworks is necessary but not sufficient to provide adequate protection. Firstly, despite improvements, national data protection legislation in a significant number of countries is inadequate, outdated, and lacking in effective enforcement. Secondly, general data protection legislation often does not apply (or apply in a limited ways) to processing of personal data for law enforcement or national security purposes.

Thirdly, AI technologies raise specific challenges for the data protection legal framework. Existing data protection laws tend to provide safeguards only in relation to the processing of personal data, i.e. data from which an individual can be identified either directly or indirectly. AI technologies often blur this distinction between personal and non-personal data. Machine learning and big data analytics, for example, are fundamentally based around the idea of extracting information from data and these technologies develop ways to identify individuals from data that would historically be considered non-personal data,

[51] See: https://privacyinternational.org/examples/quarantine-enforcement-and-covid-19 and https://privacyinternational.org/examples/tracking-global-response-covid-19
[52] See: https://towardsdatascience.com/drones-and-artificial-intelligence-to-enforce-social-isolation-during-covid-19-outbreak-783434b7dfa7
[53] PI, "10 questions to Palantir from privacy organisations", https://privacyinternational.org/press-release/3732/press-release-10-questions-palantir-privacy-organisations.
[54]See: https://privacyinternational.org/news-analysis/3461/extraordinary-powers-need-extraordinary-protections
[55] See: https://www.pewtrusts.org/en/research-and-analysis/articles/2021/03/11/artificial-intelligence-has-helped-to-guide-pandemic-response-but-requires-adequate-regulation
[56] See: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30295-8/fulltext
[57] See for further information PI, "Data Protection Guide", available at https://privacyinternational.org/data-protection-guide.

and therefore outside the purview of data protection law. AI applications may also blur the distinction between sensitive and non-sensitive personal data. Certain categories of personal data, similar to protected characteristics, are usually considered more sensitive, and are thus subject to higher protections. Through advanced data analytics, highly sensitive details revealing or predicting an individual's sexual life, health status, religious or political views, can be gained from seemingly mundane data.

Further, AI applications may rely on non-personal data to make or inform decisions that still negatively impact the human rights of individuals and groups affected. In these circumstances, data protection law offers little in ways of protection.

In assessing the adequacy of the national legal framework to protect human rights, it is therefore necessary to consider the wider range of laws relevant to AI technologies, including equality, consumer protection, electronic safety, product liability, competition, redress and administrative law, to name a few, together with sectoral legislation governing the deployment of AI applications in specific sectors, such as health care, criminal justice, immigration control, financial and insurance sector, etc.

- **Public procurement of AI applications**

Because of the increasing reliance by governments on AI applications for the delivery of a wide array of public services, PI believes that specific attention should be paid on the legislative framework governing public procurement of AI technologies and the safeguards to be put in place in contracting public services to private companies employing AI technologies. In ourresearch on the public-private surveillance partnership, PI has identified some common concerns related to: lack of transparency and accountability in the procurement processes; failure to conduct due diligence assessments; growing dependency on technology designed and/or managed by private companies, with loss of control over the AI applications themselves (to modify, update, fix vulnerabilities, etc.), over-reliance on the technical expertise of the private company and there are also risk of vendor lock-in. In many cases, the private company supplies, builds, operates and maintains the AI system they deployed, with public authorities not having sufficient knowledge or effective oversight. Lack of adequate legal framework is often compounded by limited enforcement safeguards provided for in contracts, resulting in limited or no venues for redress.[58]

5. Safeguards

There are certain specific safeguards that are key to ensuring the protection of the right to privacy when designing and deploying AI technologies and that should thus be enshrined in law, implemented and enforced.

5.1 Ensuring transparency, interpretability and explainability

The opacity of complex AI applications poses significant challenges to accountability and ultimately to access to effective remedies. However, not all sources of opacity are of a technical nature and many can be addressed by adopting a human rights centred approach. This is particularly the case when opacity is due to proprietary software and trade secrets; deliberate opacity by design; or lack of technical expertise that is required to properly understand advanced processing using AI.[59]

[58] See: https://privacyinternational.org/learn/public-private-surveillance-partnerships

[59] As noted in the Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems: "the legislative frameworks for intellectual property or trade

Data protection standards, such as the right to information, articulate some transparency requirements. Information shall include the category, purpose and sources of the data processed; the existence of profiling, of automated-decision making; and the logic involved and the significance and envisaged consequences of the processing. This may be elaborated further to include for example "factors taken into account for the decision-making process, and their respective 'weight' on an aggregate level" and how a profile was built "including any statistics used in the analysis".[60] Such an obligation should apply even where the task is burdensome.[61] The domestic legal system, including intellectual property and trade secrecy, should not preclude transparency of AI applications.

5.2 Respecting human rights by design

Decisions made in the design stage of AI application have a significant impact on whether the technology is human rights compliant. Relevant factors that would affect the design of an AI application include: deciding which processes will be automated; setting the values the AI application is designed to optimise; assessing the training data used; deciding in which circumstances the AI application shall be used.[62]

Data protection legislation often includes obligations of privacy by design, requiring inter alia to ensure that the design of AI applications which process personal data limit data collection, restrict further data processing, prevent unnecessary and unauthorised access, amongst other privacy enhancing measures. These measures should all be part of the design of AI applications, but they should be complemented by considering other measures aimed at addressing other human rights risk factors. For example, testing and evaluation of AI application should consider the specific context in which they are intended to be deployed; the data to be used in testing should allow to mitigate risks of bias and discriminatory outcomes. These requirements and safeguards should be built in laws that regulate AI technologies in the relevant sectors, for example in healthcare.

5.3 Human Rights Impact Assessment

Human rights impact assessments of AI applications should be conducted at all stages of the AI applications: prior to the design, during the development, the testing, the deployment and regularly thereafter in order to identify the emerging human rights risks.

---

secrets should not preclude such transparency, nor should States or private parties seek to exploit them for this purpose. Transparency levels should be as high as possible and proportionate to the severity of adverse human rights impacts, including ethics labels or seals for algorithmic systems to enable users to navigate between systems. The use of algorithmic systems in decision-making processes that carry high risks to human rights should be subject to particularly high standards as regards the explainability of processes and outputs." https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154

[60] Article 29 Data Protection Working Party, Guidelines on Automated Decision-Making and Profiling for the Purposes of Regulation 2016/679, 17/EN. WP 251rev.01, 6 February 2018, p 27.

[61] The Article 29 Working Party Guidance on Transparency (adopted by the European Data Protection Board) has underlined that "[...] the mere fact that a database comprising the personal data of multiple data subjects has been compiled by a data controller using more than one source is not enough to lift this requirement if it is possible (although time consuming or burdensome) to identify the source from which the personal data of individual data subjects derived. Given the requirements of data protection by design and by default, transparency mechanisms should be built into processing systems from the ground up so that all sources of personal data received into an organisation can be tracked and traced back to their source at any point in the data processing life cycle." https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227

[62] For some examples of the factors to consider see comments by Privacy Researchers on the proposals of the Office of the Privacy Commissioner of Canada (OPC) to amend the Personal Information Protection and Electronic Documents Act (PIPEDA) for ensuring appropriate regulation of artificial intelligence, https://tlpc.colorado.edu/wp-content/uploads/2020/03/2020.03.13-Academic-Researchers-Comment-on-ensuring-appropriate-regulation-of-artificial-intelligence-final-1.pdf.

These assessments not only enable the identification of the risks and corresponding mitigation strategies required to respond to them, but they also provide a framework for deciding whether to go ahead with a particular initiative. The outcomes of the assessment should result in redesign or cancellation if the risks outweigh the benefit.

While certain AI applications which carry significant risks for human rights (due to the technology used and/or the sector in which they are used, see above) require additional scrutiny, PI believes that at a minimum, an impact assessment should include privacy and data protection impact assessments as well as an assessment of other human rights likely affected by the AI application as well as potential discriminatory effects. Such assessments should consider the necessity and proportionality of any interference with privacy or other human rights, the risks to individuals and groups, and how these risks are to be addressed and mitigated.

The assessments should be conducted with the participation of affected individuals and groups, civil society actors and independent experts. The outcome of the assessment should be made public and should detailed the mitigation and oversight measures envisaged. As noted by the Committee of Ministers of the Council of Europe "confidentiality considerations or trade secrets should not inhibit the implementation of effective human rights impact assessments."[63]

5.4 Security of AI

The security of the data, at rest and in transit, as well as the infrastructure relied upon for processing, should be protected by security safeguards against risks such as unlawful or unauthorised access, use and disclosure, as well as loss, destruction, or damage of data.[64]

When assessing the level of security for AI applications, organizations should consider central processing and data storage sites, as well as the security of remote devices where data also may be collected or received. Security measures should include appropriate mechanisms for addressing actual and suspected security breaches. PI research has shown how cheap smart phones are often marketed with pre-installed apps which not only collect personal data without users' ability to control, but are also riddled with vulnerabilities which can be easily exploited, particularly because of lack of security updates.[65] As PI's correspondence with Google outlines, big tech companies have an important role to play to ensure the security and privacy of devices, including by prohibiting certain practices which put privacy and security of users' data at risk.[66]

5.5 Independent oversight

Any deployment of AI technology should be subject to independent, effective, adequately resourced and impartial oversight. Oversight should cover all parts of the design, use and throughout the deployment of AI application.

Oversight depending on the type of technology and the sector in which it is deployed should include judicial, administrative and/or parliamentary domestic oversight mechanisms capable of verifying the legality of the use of AI, ensuring transparency and

[63] Council of Europe, Addressing the impacts of Algorithms on Human Rights, Recommendation of the Committee of Ministers, https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154
[64] PI, "A Guide for Policy Engagement on Data Protection; The Keys to Data Protection", https://privacyinternational.org/sites/default/files/2018-09/Data%20Protection%20COMPLETE.pdf.
[65] See https://privacyinternational.org/long-read/3226/buying-smart-phone-cheap-privacy-might-be-price-you-have-pay
[66] See https://privacyinternational.org/news-analysis/4118/our-response-google-privacy-isnt-luxury

accountability. Oversight mechanisms should be able to verify the fairness and accuracy of AI application.

Oversight mechanisms must have the power and capacity to conduct regular auditing of AI applications to ensure their compliance with human rights and other standards. As noted by the UN Special Rapporteur on freedom of expression, protection of intellectual property and trade secrets cannot justify refusal of such oversight, particularly when the AI application is used by the public sector. Further, there are technical and policy options to address legitimate concerns related to proprietary technology, including allowing regulators and independent researchers access to AI applications on a confidential basis.[67]

5.6 Ensuring access to remedies – both individual and collective

Individuals should have access to an effective remedy against applications of AI technologies that affect them. As access to a remedy is dependent on the ability to know if and how one has been affected by AI applications, transparency and explainability noted above are necessary preconditions to exercise the right to seek remedy.

Individuals should have access to accessible, affordable, independent and effective judicial and non-judicial authorities with the power to receive complaints from individuals, investigate them, and take enforcement action - or refer the case to a court. As noted by the UN Special Rapporteur on freedom of expression, there are concerns whether AI applications, such as automatic response processes, to respond to complaints constitute an effective remedy, "given the lack of discretion, contextual analysis and independent determination built into such processes."[68]

Beyond individual redress, mechanisms of collective redress are an important and effective tool for accountability of AI applications. As noted above, challenges in transparency and explainability and the fact that AI systems often affect groups and communities, as well as the society more broadly, make collective complaints appropriate procedure to complement individual redress.[69]

[67] Report of the UN Special Rapporteur on freedom of expression, 29 August 2018, UN doc. A/73/348.
[68] See UN doc. A/73/348, para 41.
[69] See for some examples of PI's complaints: https://privacyinternational.org/legal-action/complaints