Report of the proceedings of the online expert seminar with the purpose of identifying how artificial intelligence, including profiling, automated decision-making and machine learning technologies may, without proper safeguards, affect the enjoyment of the right to privacy (27-28 May 2020)

1. Introduction

Pursuant to operative paragraph 10 of resolution 42/15 of the Human Rights Council, the Office of the United Nations High Commissioner for Human Rights organised an expert seminar to discuss how artificial intelligence (AI), including profiling, automated decision-making and machine-learning technologies may, without proper safeguards, affect the enjoyment of the right to privacy.

The seminar discussed the impacts of AI on the enjoyment of the right to privacy and the role that the right to privacy plays in safeguarding other human rights affected by AI technologies. The seminar also sought to articulate legal and regulatory frameworks and safeguards that States, businesses and international organisations are required to put in place to promote and protect the right to privacy when designing, developing, deploying and evaluating AI technologies.

Due to COVID-19 pandemic, the expert seminar was held online and webcasted live. Recordings are available at http://webtv.un.org/search?term=%22Expert+Seminar+on+Artificial+Intelligence++Right+to+Privacy%22&sort=date

2. Opening statement by the representative of the UN High Commissioner for Human Rights

Peggy Hicks, Director, Thematic Engagement, Special Procedures and Right to Development, Office of High Commissioner for Human Rights, opened the seminar by noting how AI technologies, including profiling, machine-learning and automated decision-making, increasingly permeate everyone's lives, economies, and societies. She observed that as ever more powerful analytical tools continue to emerge, they organise and interpret a vast array of seemingly uncorrelated data; they automate decision-making, cutting time and, notably, cost; they predict behaviour and events at individual and societal level.

Ms. Hicks underscored that AI technologies could be a powerful tool for advancing human progress but, if applied without safeguards, they posed significant risks to human rights and to the right to privacy in particular, as noted by the Human Rights Council in resolution 42/15. For example, in the current global COVID-19 crisis AI was indispensable for finding a vaccine, but at the same time, there was a rush towards AI-powered tools to trace and track at times without demonstrable benefits and at the cost of carrying out and normalizing mass surveillance.

Ms. Hicks remarked that AI could power pervasive surveillance, online and offline; it could be used to infer personal traits and profile and predict people's behaviour, shrinking the already small private space people may have to be themselves. She remarked that the potential impact of this profiling and targeting on democratic processes had become one of the most important and debated digital issue in recent years.

Ms. Hicks also recalled that UN experts have identified some of the implications of the use of AI on civil, political, economic and social rights. She referred to the Special Rapporteur on extreme poverty who had warned against the risks of the current trends towards digital welfare benefits, from 'unrestricted data matching" to surveillance enabling 'around the clock monitoring of beneficiaries', to imposition of conditions that 'undermine individual autonomy'.

Recalling the recent UN Secretary-General report to the Human Rights Council on the role of new technologies for the realization of economic, social and cultural rights (A/HRC/43/29), Ms. Hicks reaffirmed that international human rights law offered the framework to regulate the use of AI technologies. She added that the design, development, deployment and evaluation of AI technologies must be in line with the obligations of States under international human rights law and the responsibilities of business enterprises, as set out in the UN Guiding Principles on Business and Human Rights. These include obligations to refrain from using AI in ways that violate the right to privacy or other human rights and to ensure that any interference with the right to privacy complies with the principles of legality, necessity and proportionality.

Ms. Hicks underscored the importance of effective national legal frameworks, based on human rights law. She noted that data protection law does not offer a panacea against all potential interferences with the right to privacy and other human rights by AI technologies and that it is necessary to ensure that other national legal frameworks, including sectoral laws regulating the use of AI technologies in the delivery of public services, in criminal and immigration matters, and in relation to access to insurance and financial services, are adequate to respect and protect human rights.

Ms. Hicks added that international human rights law also required that use limits to AI are set, and red lines drawn. She noted the increased support for the banning (or at least a moratorium) on the use of facial recognition technology in public spaces, on the grounds that such technology would inevitably violate the right to privacy, as well as the right to freedom of peaceful assembly.

Moreover, beyond the law, there was a plethora of safeguards, such as human rights due diligence, privacy by design and by default, transparency and adequate explainability of AI technologies that, if applied from the outset, could address some of the human rights risks posed by AI technologies.

Ms. Hicks concluded by noting that exercising effective monitoring and oversight over AI was fundamental to ensure accountability. She pointed to the significant risk posed by AI technologies in deflecting responsibility, hiding human accountability behind automated processes. She affirmed that the effects of AI technologies over human rights, whether positive or negative, reflected the political and business choices of governments and companies, and that human rights standards and principles provided the framework to ensure that the right choices were made.

3. Session I: Setting the scene: how does artificial intelligence (AI) affect the enjoyment of the right to privacy?

This first session was devoted to a discussion of the specific risks to the right to privacy presented by AI, including profiling, automated decision-making and machine-learning technologies.

The session was chaired by Renata Avila, Executive Director, Fundación Ciudadanía Inteligente. Panellists were Joe Cannataci, UN Special Rapporteur on the right to privacy; Yves-Alexandre de

Montjoye, Assistant Professor, Imperial College; and Vidushi Marda, Digital Programme Officer, Article 19.

Remarks of panellists

The first speaker, Joe Cannataci, UN Special Rapporteur on the right to privacy, presented the working group on corporations his mandate had set up. The group was working on a set of principles on AI and the right to privacy. The Special Rapporteur noted that one of the main problems of regulating AI in law was the lack of technical knowledge of lawyers and law makers. In addition, he stressed the need for data protection laws and effective independent oversight, noting that around one third of UN member states had no privacy law and only 65 to 70 states had independent data protection authorities. He remarked that effective data protection laws also regulate AI technologies, including governing AI uses of medical and insurance data. Finally, the Special Rapporteur pointed out that oversight authorities are currently lacking technologists, because they lack the financial resources to recruit them.

Yves-Alexandre de Montjoye focused on technological approaches to safeguard privacy, and in particular on anonymization methods. He described AI as a set of fairly complex, highly optimised algorithms that learn from data. The increase of AI technologies and of their power was first and foremost related to the availability of data, a lot of which is data generated by individuals (e.g. mobile phone data, location data, internet searches and shopping habits.) He remarked that when privacy concerns are raised about the use of such data, the response is often that the data used is anonymous or depersonalised. However, de Montjoye explained, the techniques of psyeudonimisation, de-identification, and uncertainty/sampling are no longer effective with new AI technologies and the availability of large data sets (including cross referencing to publicly available data). His research suggested that there is no reason to believe that an efficient enough, yet general, anonymization method will ever exist for high-dimensional data. Anonymisation was not sufficient per se to exclude data from being considered personal data for the purpose of data protection regulation. These findings were confirmed by a range of recent examples of re-identification of anonymised data sets. He nevertheless cautioned against criminalizing re-identification of anonymised data. Instead he advocated for the use of modern privacy-preserving techniques, such as differential privacy.

Vidushi Marda noted the trend to attribute the achievement of desirable social goods to new technologies and the need to identify where these technologies are leading us. She raised concerns about experimentation and trials of machine learning technologies (sometimes at population scale, such as with Adhaar in India) which had profound impacts on the rights to freedom of expression and privacy. She cited as an example the deployment of facial recognition technologies without data protection safeguards. She argued that human rights assessment needed to be conducted before the technology is deployed, at the design/conceptualisation phase. Ms. Marda also advocated for a shared language between laws, ethics, human rights and technologies. She argued that we need technology that is legally informed but also laws that are technically informed. Noting the usefulness but also the limits of ethical guidelines and frameworks, she stressed the need to develop human rights-based approaches to AI. Concluding on the issue of transparency, she opined that transparency as a technical solution (e.g. access to the source code) was often not enough. Instead, transparency needed to be embedded in the process of decision making, design and development of AI technologies. Excluding human rights concerns from consideration at the early stage of the design and conceptualisation of AI leads to human rights issues that are not easily addressable at the deployment stage.

Summary of the discussion

Responding to the question of which issue should be prioritised in this field, the Special Rapporteur on privacy recommended that any AI technologies should be deployed only after conducting strict privacy impact assessments. Ms. Marda suggested that it was necessary to debate whether a particular technology should be deployed in the first place, assessing whether it was fit to tackle a specific problem, rather than being presented as a solution without questioning it. She expressed particular concerns at the risks of decision making by technologies that are increasingly monopolised by companies and governments and/or by few states. Yves-Alexandre de Montjoye raised concerns at the false dichotomy being presented between having competitive AI technologies and protecting privacy, including applying data protection rules. With the right technical and policy tools, it was possible to have AI and preserve privacy.

Prof. Cannataci and Prof. de Montjoye noted that privacy engineering should be taught and understood more widely. Ms. Marda urged all stakeholders to hear the testimonies of those affected by technologies and organisations, including grass-roots organisations, which represent them. Ms. Avila added that feminist and gender perspectives had to play a central role in discussions and decision-making around AI.

Intervening in the discussion, Eduardo Bertoni, Director, Argentina National Access to Public Information Agency, reaffirmed the need to ensure that lawyers and policy makers had the technical understanding of the technologies they seek to regulate and of the capacity and limits of technologies in addressing social needs, including in the context of the COVID-19 pandemic.

During the ensuing discussion, the representative of the Permanent Mission of Switzerland noted the importance of human rights and users-based approach to AI. He recalled the responsibilities of the private sector under the UN Guiding Principles on Business and Human Rights. He stated that interferences with privacy affected other human rights; social media surveillance and the use of facial recognition technologies, for example, affected the rights to freedom of expression, of peaceful assembly and association. He stressed the need for free and informed consent and purpose limitation, legality, necessity and proportionality, human rights due diligence and impact assessment on a regular basis, access to remedy in the design of AI technologies. The representative of the Permanent Mission of India noted the importance of the right to privacy in the context of AI technologies and emphasized that AI had significant impact for developing societies such as India.

The National Human Rights Commission of India noted the enormous potential for public goods, but also the human rights challenges posed by AI technologies, most notably in relation to data mining techniques. She recalled the judgment of the Supreme Court of India of 2017 recognising the right to privacy as protected under the Indian constitution and noted the need to establish data protection framework as well as sectoral regulatory frameworks to regulate AI technologies as well as encouraging AI developers to apply international standards protecting privacy.

In his concluding remarks the UN Special Rapporteur on the right to privacy noted that AI technologies had been around for decades. In Europe there were significant legislation and policy developments including model legislation that could be used. He deplored that political will is often missing among UN member states, particularly within the Human Rights Council. Ms. Marda concluded by recommending that accuracy and efficiency needed to be accompanied by transparency and accountability, challenging the notion of using AI technologies that are inscrutable. Further, she stated the need to apply existing laws, and principles, including human rights, data protection, consumer protection, to AI technologies, before introducing new laws. Prof.

de Montjoye reaffirmed the need to develop and apply safeguards to ensure that AI technologies be developed and used in ways that respect the right to privacy.

4. Session II: The right to privacy and the protection of other rights affected by AI

The second session focussed on the importance of upholding the right to privacy and on how privacy interference by AI applications (by governments and business enterprises) can undermine other human rights and principles, such as the right to peaceful assembly, freedom of expression, access to health, social security and non-discrimination.

The session was chaired by Sophie Kwasny, Head of Data Protection Unit at the Council of Europe. Panellists were Lorna McGregor, Director, Human Rights, Big Data and Technology Project, Essex Law School; Amos Toh, Senior Researcher, Artificial Intelligence and Human Rights, Human Rights Watch; and Chenai Chair, Web Foundation Gender and Digital Rights Research Manager and Mozilla Tech Policy Fellow.

Remarks of panellists

Lorna McGregor noted that the right to privacy risks being traded-off, if the recognised principles of legality, necessity and proportionality are not applied to assess privacy interferences.

By way of illustrating these risks, she elaborated on the deployment of contact tracing apps in the context of the COVID-19 pandemic. Earlier proponents argued that the right to privacy needed to give way to allow states to fulfil their obligation to protect health and lives. Framing this issue in dichotomy/oppositional terms (or as a 'trade off'), however, puts the right to privacy at risk. Contact tracing apps could lead to serious interferences with the right to privacy, depending on a range of factors (such as what data is collected, where it is stored (centrally or locally), for how long, which agencies or companies may have access to the data now or in the future. Further, the adoption of this technology could be a gateway for future use for other objectives, potentially leading to a surveillance state. Prof. McGregor stressed that the right to privacy acted as a gatekeeper for other human rights, and that articulating the effects on privacy of technologies such as the contact-tracing apps could help clarifying the impact on other human rights.

Prof. McGregor argued that international human rights law provided the framework for states and businesses to manage the relationship between the right to privacy and other human rights or interests and to achieve the dual objective of protecting privacy and health. The human rights framework helped narrowing the technological options available, thereby limiting the impact on privacy. It also required transparency in the adoption and application of these technologies and public debate and scrutiny, including on the scientific justification of the apps and on its place in the overall strategy to combat the pandemic. Human rights law required a clear and accessible legal basis, and also that States demonstrate the necessity and proportionality of the measures taken. Oversight and accountability structures must be established and full human rights impact assessment carried out. Prof. McGregor concluded that legislation was emerging which, while not perfect, aimed at honouring the human rights framework, with privacy protected as well as public health and other rights.

Amos Toh addressed the interplay between the right to privacy and the right to social security. He pointed to trends of making access to welfare increasingly conditioned or regulated by automated systems. Recalling its early application, Mr. Toh noted that the Welfare Reform of 1996 in the USA

introduced fingerprint verification as one of the antifraud measures. He noted, providing examples from Ireland and the UK, that wrapped into the seemingly legitimate aim of fraud detection and anti-corruption measures was the unspoken presumption that those seeking to access welfare benefit were seeking to cheat the system. He recalled the decision of the Hague district court in the Netherlands to strike down an automated fraud detection tool as a violation of the right to privacy due to the opacity of the program and the way risk scores were attributed.

Chenai Chair noted that the discrimination resulting from the application of AI technologies was an effect of the existing discrimination in our societies. She remarked that the race to AI in developing countries was mostly focussing on the economic benefits and did not reflect upon existing gender and digital divides. Recalling examples from recent research, she stressed that discrimination resulting from AI was mostly related to gender, sexuality, race and social economic status. This research pointed towards the need to query how these AI technologies are designed and how to ensure not only privacy by design but also non-discrimination by design. Impact assessment had to look into the societal and potential discriminatory effects and not only the economic gains of AI.

Summary of the discussion

Responding to a question on oversight mechanisms, Prof. McGregor noted that effective oversight structures depended on the technologies under scrutiny and the powers available to existing mechanisms. She acknowledged the important role of internal ethical boards, but also underscored that international human rights law required independent oversight mechanisms. Data protection authorities were necessary but might not be sufficient to monitor the impact of AI on the full range of human rights.

Reacting to remarks from the audience, Ms. Chair noted that there was little awareness of how to exercise the right to privacy online, particularly when faced with obscure terms of service and practices (such as opt in options) that do not embed privacy by design in the digital communication context.

Mr. Toh reiterated that ‘trade off’ is a very dangerous framing affecting the rights to privacy and to health in the context of the responses to COVID-19. He also argued that some of the AI technologies being introduced raise significant discriminatory and inequality concerns, related to digital divide but also to a disconnect between assumptions underpinning mobile tracking technologies and the living reality of many people (such as refugee communities who often share same mobile phones, making individual tracking ineffective.)

A representative of Derechos Digitales raised concerns about the use of personal data to deliver services in Latin American countries, particularly in contexts where reliance on privately owned infrastructures or services resulted in personal data being transferred to commercial entities. These projects were forcing the population to surrender the right to privacy to access other fundamental rights. In response, Prof. McGregor noted that the international human rights framework needed to apply to data sharing practices between states and companies, and between companies, and called for effective oversight of these practices.

In response to a question on the link between the right to privacy and the right to freedom of information, Mr. Toh noted that these rights were closely interconnected not only in the context of content filtering and moderation but also on accessing information on social media platforms, which uses algorithms trained on personal data. The desire for targeting content required data collection on users. Most content moderation and content delivery practices were still largely opaque, leaving

little information on how and why ads or other content are delivered to individuals online. Building on this, Prof. McGregor noted that it was increasingly clear that protecting the right to privacy was at the core of addressing targeting practices online. In that context, it was necessary to address the limits of relying on individual's consent as a legal basis for processing personal data, given how limited control and understanding individuals have on how their data is used in digital contexts.

A representative of Red de Defensa de los Derechos Digitales raised the issue that AI technologies may undermine the presumption of innocence and due process standards. For example, the use facial recognition for surveillance purposes could lead to an innocent person being accused of a crime; automated sentencing was another worrisome possibility. Prof. McGregor noted the difficulties to challenge judicial decisions made with the support of AI technologies and the risks involved in the judicial actor deferring to AI without understanding how the technology operates, for example in producing a risk score on an individual. Ms. Chair also noted the need to qualitative research to assess how these automated processes are supporting judicial decisions and whether they are free from discriminatory effects.

The discussion's focus then shifted to regulating private sector actions. A representative of Partnership on AI raised concerns about the inadequate regulation of businesses and asked to ensure that the UN Guiding Principles on Business and Human Rights can effectively apply. This should include follow up with companies that were non-compliant with human rights impact assessments. Prof. McGregor noted the need to continue operationalising the UN Guiding Principles. She underscored that human rights impact assessments at the conception, design and deployment states of AI technologies could help internal oversight mechanisms to identify and remedy negative human rights impacts. Mr. Toh added that the UN Guiding Principles were necessary but not sufficient, particularly in the context of public services where the private sector is providing government agencies with the systems to provide public services. Transparency and access to information on development, procurement and deployment needed strong regulation rather than voluntary systems and needed to be able to override trade secrecy and commercial interest exemptions.

The Inspector General of Portugal raised questions related to predictive policing. He noted that the business case for predictive policing is predicated on the need to maximise limited resources. He highlighted four major risks: increase racial profiling, interference with privacy; over-reliance on technologies; and the risk of targeting people based on statistical probability rather than individual suspicion of wrongdoing.

Prof. Mc Gregor stressed the need to understand when these tools are used and the rationale for their deployment. She raised concerns that existing discrimination can be amplified by the introduction of predictive policing, as research of over-policing of particular communities has proven. Further indirect human rights implications needed to be assessed, including for example the ability of police to maintain and support community relations, when relying on predictive, automated policing. Mr. Toh added that predictive policing relies on historical data with embedded patterns of race, religious and other forms of discrimination. These technologies were not only fraught with risks of discrimination, but distracted attention and funding from addressing social root causes of crimes.

5. Session III: Preserving the right to privacy: legal and regulatory approaches

The third session addressed the regulatory measures necessary to address and remedy negative human rights impacts of AI technologies. The participants discussed data protection laws and other legal frameworks and how to ensure that such technologies promote the enjoyment of human rights, for example by strengthening emerging national AI strategies.

The session was chaired by Lorna McGregor, Director, Human Rights Centre, Essex Law School and Human Rights Centre. Panellists were Sophie Kwasny, Head of Data Protection Unit, Council of Europe; Eduardo Bertoni, Director, Argentina National Access to Public Information Agency; Teki Akuetteh Falconer, Director, Africa Digital Rights Hub; and Fanny Hidvégi, European Policy Manager, Access Now, Brussels.

Remarks of panellists

Teki Akuetteh Falconer focussed her remarks on developing countries, such as on the African continent, noting the challenges of regulating technologies. She pointed out that significant challenges existed. At regional level, the African Union Convention on Cybersecurity and Data Protection had not entered into force because of lack of ratification. More than 25 countries in Africa did not have data protection laws. Only around half of those which do have a functioning data protection authority in place. Ms. Akuetteh Falconer noted that due to lack of resources and capacity, national data protection authorities had yet to adopt plans or guidelines to address the challenges that AI technologies pose to data protection. Similar questions about resources and capacity of regulators arose in the context of possible AI specific legal frameworks at national level. She suggested using the existing data protection framework to engage all stakeholders in developing standards and principles aimed at human-centred AI, based on dignity, liberty and freedom for all.

Eduardo Bertoni focussed on the Latin American context. In 2019, the Ibero-American Data Protection Network published General Recommendations for the Protection of Personal Data in Artificial Intelligence, providing recommendations to developers and manufacturers of AI. He illustrated the ten recommendations developed in the document, namely (1) compliance with data protection laws, (2) carrying out a Privacy Impact Assessment, (3) embedding privacy, ethic, and security by design and by default, (4) operationalising the principle of accountability; (5) designing appropriate governance structures in organizations developing AI products; (6) adopting measures to guarantee the observance of data protection principles; (7) respecting the data subject rights and implementing effective mechanisms for the exercise of those rights; (8) ensuring data quality; (9) using anonymization tools; (10) increasing data subjects' trust and transparency (including transparency of algorithms). The Network had also released Specific Guidance for Compliance with the Principles and Rights that Govern the Protection of Personal Data in AI projects.

At the beginning of her intervention, Sophie Kwasny noted that AI was not new, although there was an unprecedented rise of its use. She affirmed that the broad definitions of personal data and of processing contained in Convention 108, modernised Convention 108, and the EU General Data Protection Regulation (GDPR) allow to address the challenges posed by new technologies, including in relation to the limits of anonymised data. Noting that Convention 108 was the only data protection treaty open to be acceded by any country of the world, she remarked that the adoption of modernised Convention 108 represented a new global benchmark. It included a number of provisions to respond to the increase in algorithmic decision making (including on data minimisation, transparency, accountability, privacy by design, data protection impact assessments,

right not to be subject to a decision based solely on automated processing, right to know the reasoning of the processing) and required stronger powers and resources for data protection authorities. She presented the specific guidelines developed by the Council of Europe Committee of Convention 108 on AI and data protection, aimed at designers, manufacturers and AI service providers, as well as legislators and policy makers; and the Recommendation of the Council of Europe Committee of Ministers on human rights impacts of algorithm systems (2020/1) addressing all human rights impact of AI technologies with guidance on obligations of states and responsibilities of companies. She also highlighted that the Council of Europe Ad Hoc Committee on AI (CAHAI) is exploring a possible legal framework on AI. Ms. Kwasny dedicated the next part of her intervention to the issue of "red lines". She pointed out that CAHAI was tasked to identify possible areas where AI should not be used, taking into account legitimacy of the technology and whether it creates a strong asymmetry between the individual and the entity using it. She also referred to the work of the Council of Europe Committee of Convention 108, which among its key priorities had identified possible red lines on facial recognition and use of AI in education systems. Against the backdrop of increased interests of political parties and other actors on the use of AI technologies, the Committee was also looking at the implication of AI on political campaigns and elections. Concluding, Ms. Kwasny supported the call of the UN Special Rapporteur on right to privacy to adopt data protection legislation, and to join the modernised Convention 108.

Fanny Hidvégi noted that drafting laws to cover AI faced difficulties to define their material scope. More fundamentally, the problem legislation on AI needed to address was not regulating a particular technology, such as machine learning, but rather the context of human rights violations (at individual and collective level) and how these might be enabled by new technologies. She noted that beyond the debate on a general regulation of AI, there were sector specific applications of automated technologies that were already regulated into laws. She also expressed scepticism about the concept of regulatory 'sandboxes', describing it as a most contradictory approach from a human rights perspective. She stated that a law with AI in its title was not technology neutral and future proof; however she recognised the need to enforce human rights horizontally and to introduce new safeguards.

She then turned to Access Now's contribution to the European Commission's white paper on AI, whose recommendations can be applied beyond the EU specific context. To start with, she recommended to not indiscriminately promote the introduction of AI technologies without first a demonstrable societal benefit. Secondly, a rights-based approach needed to be implemented and human rights impact assessments (including due diligence) conducted throughout the life cycle of AI systems. Thirdly, biometric technologies that enable mass surveillance, should be banned. Fourthly, national centres of AI expertise, supporting existing regulators (including human rights), should be established, and an equality body should decide on discrimination cases. Lastly, she recommended to enforce high scientific standards, include public interest criteria for research and funding, and disregard pseudoscience, such as emotional recognition and prediction technologies which lack scientific basis.

She noted that the Freedom Online Coalition led by Canada established a task force on AI to ensure respect of human rights. The prevention of violations by AI technologies required systematic reform of human rights enforcement. Concluding, she noted that new safeguards, based on international human rights law, must be in place to prevent and mitigate new opportunities of government and private actors to abuse their powers with AI technologies.

Summary of the discussion

Launching a lively discussion, Ms. Falconer noted that courts had an important role to play, particularly in politically polarised environments where courts become the last resort for the enforcement of law. However, judges and lawyers often had limited understanding of the technologies. This issue could be addressed by the establishment of specialised courts. Further she noted the need to develop common industry standards which would help minimise the challenges of enforcing laws across different jurisdictions for multinational private actors, including technology giants.

Mr. Bertoni argued that legislation needed to be technology neutral. Otherwise the law would need to continuously catch up with the development of the new technologies. He noted that privacy impact assessments still faced implementation challenges and reminded the participants of the long time it took to develop environmental impact assessments. He added that building on the UN Guiding Principles on Business and Human Rights, there was a need to continue engaging with private actors, to ensure they accept and meet their responsibility to respect human rights.

Responding to a question by a representative of the Permanent Mission of Germany on how to ensure compatibility of existing and new regulations on AI, Ms. Kwasny noted that the Council of Europe invited all stakeholders to contribute to ensure a broad understanding of the proposals being discussed. She noted that any instruments developed will be in the framework of the founding principles of the Council of Europe, fundamental human rights, democracy and the rule of law, and might address questions of ex-ante control, liability and discrimination which are not specifically addressed by the data protection standards.

A representative of Privacy International agreed that data protection laws were necessary but not sufficient. She noted that it was necessary to take into account specific applications of AI technologies in certain situations of heightened risk of human rights violations, giving the example of deployment of AI-supported lie detectors for border management.

Ms. Hidvégi underscored a practical problem that civil society organisation faced. Giving their limited resources, they were unable to effectively follow multiple negotiations of sectoral legislation. She proposed adopting a horizontal approach on safeguards that are missing in the human rights toolbox, for example mandatory human rights impact assessments. She also noted the need to make the business case for compliance and promotion of human rights, and added that some regional players, like the European Union, could leverage their economic power to ensure support to human rights approaches to AI.

Responding to a question by the University of Istanbul, Ms. Falconer noted that consent was one of several legal grounds for processing personal data. She also highlighted the challenges of using consent as legal ground for processing of personal data by AI technologies. In this regard, certain sensitive categories of data would require consent and more stringent safeguards and oversight.

A representative of the Permanent Mission of Maldives inquired about existing or future model legislation. In response, Mr. Bertoni noted that international or regional treaties, like Convention 108, as well as national laws provided the general principles and framework. They were often accompanied by interpretation and guidelines for specific technologies by international or national monitoring authorities.

Summarizing the session, Prof. McGregor underscored that a multi-layered approach to regulating AI technologies was necessary. She highlighted the core importance of effective data protection law as well as the critical role of data protection authorities, human rights and equality bodies.

6. Session IV: Preserving the right to privacy: due diligence, data governance and other safeguards

This session discussed a range of procedural steps and technological solutions that States and businesses should take to prevent violations of the right to privacy when using AI, and how to address and remedy them where they occur. A particular focus was on human rights due diligence approaches and technical solutions.

The session was chaired by Mila Romanoff, Privacy Officer, UN Global Pulse. Panellists were Fionnuala Ní Aoláin, UN Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism; Rachad Alao, Engineering Director, Facebook; and Isabel Ebert, Researcher, Institute for Business Ethics, University St Gallen.

Remarks of panellists

Fionnuala Ní Aoláin explained how biometric data and AI-based tools were increasingly used in areas relevant to her mandate, particularly in law enforcement, intelligence gathering, and border management. According to her, the normalisation of the use of these technologies began with UN Security Council resolution 1373 which sought to strengthen border control and culminated with UN resolution 2396, adopted under Chapter VII of the UN Charter, which imposes legal obligations on UN member states, including to 'develop and implement systems to collect biometric data'. She noted that due to State practices, counter-terrorism measures applied to vast sectors of populations, including human rights defenders, religious minorities and others. In this context, she pointed out that the proliferation, use and normalisation of advanced technologies in counter-terrorism was not accompanied by human rights guidance, particularly in relation to biometric technologies. She raised particular concerns at the trend, identified by her mandate, of subcontracting the regulation and development of guidance on counter-terrorism measures to a range of new institutions with unclear legal standing, selective membership, and lack of human rights mandate or expertise.

She illustrated the findings of her forthcoming report on human rights compliant use of biometric technologies. This report identified a trend of widespread use of biometric technologies for counter-terrorism and highlighted particular risks when biometric technologies are employed by authoritarian states and in vulnerable situations (such as conflict zones). She acknowledged the importance of data protection laws, but also pointed out that many data protection laws had opt-out clauses for national security purposes (including the Council of Europe Convention 108).

Moreover, data protection laws were insufficient to address the overall effects of AI technologies on human rights. She also noted that the international human rights framework is adequate but its implementation is significantly deficient. There was a protection gap in relation to business enterprises, particularly in relation to deployment, selling and transfer of biometrics tools to States with poor human rights records.

Prof. Ní Aoláin outlined a range of recommendations to States, including the imperative to conduct human rights impact assessments, and that data intensive systems should only be deployed if demonstrably necessary and proportionate to achieve the legitimate aim, particularly when

deploying centralised and integrated systems. She also recommended setting up licencing systems governing technologies that represent high risks to human rights. She further listed recommendations aimed at business enterprises, including the adoption of public policy commitment to their responsibility to respect human rights and carrying out human rights due diligence. Lastly, she recommended that the UN support the development of human rights guidance on the development and deployment of biometric tools and facilitate the establishment of an international framework to govern transfer of biometric technologies.

Rachad Alao observed that Facebook had been at the centre of many news stories for its practices related to privacy. He acknowledged that the company privacy standards were not always as high as they could have been, but pointed out that in recent years it has invested significantly in privacy technologies. Facebook's application of responsible AI included a multi-disciplinary team to help the company ensure that the thousands of models that make billions of predictions everyday work in ways that are ethical and protect the privacy of users. These included the development of privacy preserving AI technology to build privacy by design into Facebook's systems. Data anonymisation and the use of differential privacy to train AI networks without collecting users' data were of particular importance. Mr. Alao noted that investment in and application of privacy technologies were not possible for companies without the relevant financial and professional expertise; therefore regulation of this space should be mindful of the ability of all business enterprises to comply. Compliance with regulations would be easier for larger entities, as it has been seen with compliance with GDPR.

Responding on a follow-up question on technological ways to de-identify facial recognition data, he explained that facial recognition technologies comprised a multistage system. It included the steps of detection, signature generation (unique representation of a face, very specific to AI system in use and difficult to make it interoperable) and matching of the unique signature with other signatures. To address some of the vulnerabilities of this technology (such as access to the unique signature, which is a form of personal data) 'data poisoning' methods were deployed to obfuscate the unique signature. Differential privacy could also be useful in certain phases to inject statistical noise to make it hard to extract personal data.

Isabel Ebert argued that 'data universalism' often adopted by many States and business was blind to cultural and political context. Each AI system was placed in specific real contexts, which often affected its functioning in ways not predicted in the testing/lab environment. She noted that AI was currently used in many sectors, including for the monitoring of work places, the administration of insurance and credit, targeted advertising, and in public/private partnerships, including for law enforcement purposes. She noted that affordability and availability continued to rise; as a consequence, AI applications were increasingly used throughout the business world, including by medium size companies.

Focusing on workplace surveillance, she observed that some analytical tools were marketed for their presumed objectivity to address biases (e.g. in recruitment to fight gender and racial biases) while they had resulted in discrimination due to their reliance on biased data. She noted that the surveillance tools deployed in the workplace (to monitor e-mails, conversations in video calls, physical meetings, emotional recognition analysis, facial recognition), often without the knowledge of the employee, compounded the intrusion on privacy by employers, who already hold personal data of their employees.

Focusing on the application of human rights due diligence to AI, Ms. Ebert presented some key recommendations, including considering the whole data life cycle, conducting data testing, ensuring

the presence of a feedback loop to remedy unintended consequences as well as opt out possibilities. She noted that differential privacy techniques might not be possible in certain AI applications when you need to know to whom automated decisions are applied and how (e.g. for assessing credit worthiness). It was hence necessary to tailor privacy and human rights-compliant technologies to specific business.

Summary of the discussion

Responding to a question, the Special Rapporteur on counter-terrorism and human rights noted how exceptionality often drives regulation, while the assumption should be to use ordinary methods until they are proven not to work. Elaborating on the responsibility of business enterprises, human rights impact assessments needed to look at the potential direct and indirect impact on human rights. She also underscored the importance of public internal accountability mechanisms particularly when companies become subcontractors of governments' activities. In this respect, she noted the trends of governments subcontracting activities to companies in areas where they face constraints under international human rights law, such as in the content moderation and biometric spheres. She gave the example of Facebook, which had its own definition of terrorism that did not conform with international law.

She reminded the participants that corporate responsibility under the UN Guiding Principles was independent of state obligations and existed over and above compliance with national law, thereby constituting an independent set of duties for business enterprises to comply with particularly when national laws are ineffective or repressive. Human rights should be embedded in AI systems by design and not be left as an afterthought.

Mr. Alao noted that large corporations were putting internal accountability and governance mechanisms in place to comply with human rights principles. Companies were subject to the legislation where they operate and had to comply with the national laws. As a consequence, they were facing challenges to meet different regulatory demands across jurisdictions.

Ms. Ebert underscored that many companies, which are not on the spotlight for their practices, still have a significant human rights footprint. Referring to the practice of data disclosure requests from government, she noted that while some large companies were beginning to question some of these requests, that was not the case for many other medium sized tech companies.

Responding to questions from representatives of Derechos Digitales and of the University of Groningen, the UN Special Rapporteur on counter-terrorism and human rights stated that there was a need to map the scope of regulation (or lack thereof) and of cooperation among different actors in the counter-terrorism sphere. Cooperation should not be treated as a human rights free zone to outsource activities that raise human rights concerns. As an example, she referred to the work of IOM on border management using biometric tools in twenty countries with little information available on their human rights impact. She also deplored the absence of an international framework which could help clarify the obligations of States when technologies are used by international entities.

According to Mr. Alao, some emerging initiatives on responsible and fair AI sought to address the potential impacts of AI on groups rather than limiting the focus on individuals, looking at fairness, equity and inclusivity of AI. Responding to a question by a representative of Derechos Digitales, Ms. Ebert noted that some steps had been taken to communicate the human rights risks of technologies to investors.

In her concluding remarks the Special Rapporteur warned that the security space was encroaching on many aspects of society, particularly in proposing technical tools as solutions to address an expanding range of issues in very complex contexts, with lack of human rights-based regulation and enforcement. The lack of resources and marginalisation of human rights in the counter-terrorism space meant that when the technical security tools are exported they raise significant human rights concerns.

Wrapping up the session, Ms. Romanoff underscored that human rights impact assessments and due diligence needed to be applied on AI tools in all cases. She encouraged businesses to put commitments to human rights into public policies. She noted that technical solutions existed but there also needed to be awareness of their limitations, such as in the context of re-identification.

7. Conclusions and recommendations

Throughout the expert seminar some common themes and trends emerged.

While there was overall recognition of the potential for AI to contribute to public good, experts warned against the trend of hyping the capacity of AI, and of adopting new technologies without assessing upfront their purposes and their capacity to achieve it.

Experts noted that while AI was not new, its reliance on data and its predictive abilities raise specific risks to the right to privacy. Infringements of the right to privacy often lead to, facilitate or contribute to violations and/or produce chilling effects on the enjoyment of other human rights. Hence experts warned of the risk of introducing false dichotomies or 'trade offs' between the right to privacy and other human rights when seeking to justify the adoption of AI technologies that interfere with privacy.

Further, experts warned that certain applications of AI technologies were likely to result in violation of the right to privacy and other human rights, notably the opaque and unregulated processing of biometric data in the context of counter-terrorism, law enforcement (including predictive policing), and border management, and they recommended considering the banning of certain AI technologies, such as facial recognition in public spaces or in schools.

There was agreement that international human rights law must continue to apply to AI technologies and in particular that any interference with the right to privacy must comply with the overarching principles of legality, necessity and proportionality.

In terms of regulation, experts recommended that States adopt or review data protection laws, ensuring that laws address the increased reliance on algorithmic decision making. This included, for example, incorporating the principles of data minimisation, transparency, accountability, the requirement of privacy by design and by default, and recognising among others the rights not to be subject to a solely automated processing based decision, and the right to know the reasoning of the processing. They warned against the abuse of exceptions and derogation clauses in data protection law, which could result in a regulatory vacuum for some of the most privacy invasive applications, including by governments.

Experts also agreed that data protection frameworks needed to be complemented with other national regulations to address the effects of AI technologies on other human rights and to ensure AI did not

produce discriminatory results. While scepticism was expressed about attempts to regulate AI with an all-encompassing law, it was agreed that sectoral laws were needed, for example to regulate the procurement and use of AI technologies by public bodies.

Experts stressed that AI's capacity to process and analyse vast amount of data challenged the effectiveness of some traditional technical safeguards to preserve privacy (such as anonymisation), and recommended the deployment of more effective tools, such as differential privacy and other privacy enhancing technologies. Conducting genuine, regular human rights impact assessments at the points of conception, design, testing, implementation and review of AI technologies was recognised as a key safeguard. Further, experts noted the need to support research into privacy-enabling technologies, resulting in concrete application of the principle of privacy by design.

There was agreement of multifaceted transparency requirements, from explainability of AI technologies (such as the logic of the automated decision-making or profiling), to the ability of independent regulators to audit algorithms; from obligations of governments to consult with experts, civil society and the public before adopting AI technologies that affect human rights, to the requirements of disclosing the use of AI technologies. Particular concerns were expressed about the opaqueness in the use of biometric technologies by law enforcement and other security agencies as well as the lack of regulation on the procurement, sale, export and transfer of such technologies.

Reflecting on the challenges to accountability and access to remedies, experts recommended to establish oversight and judicial enforcement mechanisms, such as data protection authorities, national human rights institutions, equality bodies, ombudspersons, as well as courts. They needed to be adequately resourced to monitor the effects of AI technologies on human rights, and have access to AI experts to support their mandates.