Ad Hoc Committee on Artificial Intelligence ('CAHAI'): contributions to the draft feasibility study on 'Human Rights Due Diligence for Artificial Intelligence'

Maria Pia Sacco
Chair of the IBA Working Group on AI

Working Group Members

Anurag Bana
Senior Legal Advisor, Legal Policy and Research Unit, IBA

Dr Theodora Christou
Queen Mary University School of Law

Maria Pia Sacco
Senior Legal Advisor Legal Policy and Research Unit, IBA

Prof Martijn Scheltema
Member of the Advisory Panel, Business Human Rights Committee, IBA

Sara Carnegie
Director of Legal Projects, IBA

Elise Groulx Diggs
Co-Chair, Business Human Rights Committee, IBA

Kevin O'Callaghan
Co-Chair, Business Human Rights Committee, IBA

the global voice of the legal profession®

Introduction

The International Bar Association, founded in 1947, is the world's leading organisation of international legal practitioners, bar associations, law firms and law societies. The IBA has over 80,000 members, including over 200 of the top legal firms in the world and corporate members from a diverse range of international companies. The IBA membership presently spans over 170 countries with 195 individual associations. The work undertaken by these organisations covers all areas of substantive law in addition to broader legal issues and ethics.

We are submitting our comments on behalf of the IBA's Legal Policy and Research Unit (LPRU) and the Business Human Rights Committee (BHR Committee). These, the LPRU and the BHR Committee, formed a Working Group to publish a guidance document on '*Human Rights Due Diligence for Artificial Intelligence*'.

The LPRU undertakes research projects and develops initiatives that are relevant to the rule of law, the legal profession and the broader global community. The LPRU engages with legal professionals, law firms, law societies and bar associations, governments, non-governmental organisations and international institutions to ensure innovative, collaborative and effective outcomes.

The BHR Committee aims to create awareness among lawyers in all fields of practice, of business and human rights, corporate sustainability, and more broadly ESG (environmental, social and governance (eg, conflict minerals and modern slavery transparency) principles. It works to promote the development of legal skills required to advise clients and to support law firm management in the emerging area of law relating to business and human rights, and to facilitate education and dialogue among lawyers who practice business and human rights.

The comments made in this discussion paper are the personal opinions of the Working Group members and should not be taken as representing the views of their firms, employers or any other person or body of persons (apart from the IBA's BHR Committee and LPRU) of which they are a member.

The IBA and Business and Human Rights

The *United Nations Guiding Principles on Business and Human Rights* (UNGPs),[1] unanimously endorsed by the UN Human Rights Council in 2011, represent a landmark contribution to the global debate on business and human rights.

Since 2013, the IBA has supported lawyers and bar associations in their knowledge and understanding of the UNGPs and their impact on activity for lawyers and business associations as business entities, as well as providers of legal services to other businesses. In particular, the LPRU has developed guidance documents and training tools to bridge the knowledge gap and build the capacity of lawyers to advise businesses on business and human rights related issues, including through building understanding of the UNGPs.[2]

---

1 Office of the United Nations High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights*, implementing the United Nations '*Protect, Respect and Remedy Framework*', 2011.

2 These include the *IBA Practical Guide on Business and Human Rights for Business Lawyers,* the Reference Annex (28 May 2016) and the *IBA Business and Human Rights Guidance for Bar Associations* (8 October 2015). The IBA's publications and activities on business and human rights are available at: www.ibanet.org/LPRU/Business-and-Human-Rights-for-the-Legal-Profession.aspx.

The IBA's work on artificial intelligence (AI)

In September 2019, the IBA published a report on '*The Future of Work*'.[3] This document is the result of a three-year collaboration between the IBA and the International Labour Organization (ILO) and focused on the legal dimensions and implications of disruptive technologies (including AI).

In June 2020, the IBA published a report on '*Digital contact tracing for the Covid-19 epidemic: a business and human rights perspective*', in which the human rights risks associated with contact tracing apps are highlighted and the role of the private sector, in preventing and addressing these risks, is discussed.

The IBA supports the values upon which the Council of Europe is founded and which have human rights, democracy and the rule of law at its centre. In August 2019, the Working Group submitted its written comments on the draft recommendation of the Council of Europe Committee of Ministers to member states on the human rights impacts of algorithmic systems. This submission was part of the IBA's work on BHR and aimed to shed light on the legal challenges faced by business entities, when preventing, mitigating and addressing the adverse human rights impact associated with artificial intelligence. In June 2020, the Working Group was admitted as an observer organisation to the Council of Europe Ad Hoc Committee on Artificial Intelligence (CAHAI).

Turning to the CAHAI's call for contributions to the draft feasibility study of 16 June 2020 ('Draft Feasibility Study'), the Working Group's suggestions and thoughts are set out in this discussion paper below. We have addressed the issues thematically and do not seek to address each and every question listed under the ten headings identified by the Draft Feasibility Study.

CAHAI Draft Feasibility Study 16 June 2020: comments

'3 – Opportunities and risks arising from the design, development and application of artificial intelligence on human rights, the rule of law, and democracy. 'Green' and 'red' areas –meaning respectively positive and problematic examples of artificial intelligence applications from a human rights, rule of law and democracy perspectives, while considering the context-sensitive environment for artificial intelligence design, development and application in Europe and developments at a global level.'

a) AI is changing society and is deployed in very diverse sectors, such as banking, human resources, risk management, law enforcement, humanitarian assistance, agriculture and medical diagnostic. Notwithstanding its disruptive impact on our lives, there is no universally accepted definition of AI.[4] For the purpose of this discussion paper and, consistently with *CAHAI (2020) 06-fin*, the IBA Working Group will refer to the definition provided by the European Commission High-Level Expert Group on AI in 2019:[5]

> 'Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems

3 *IBA Report on the Future of Work: Special Consideration to Law and Disruptive Technologies,* September 2019, available at: www.ibanet.org/Document/Default.aspx?DocumentUid=7AE9FC31-ABF4-49D7-974F-F246031825AE, last accessed 6 September 2020.

4 *CAHAI (2020) 06-fin.*

5 EU High Level Expert Group on AI, A definition of AI, main capabilities and scientific disciplines, 2019.

can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.'

From the definition above, at least, two typologies can be identified:[6] knowledge-based systems and machine learning and deep learning. The former is based on a closed, top-down approach, in which a specific set of axioms are used to support the reasoning in a given domain. The latter is based on a bottom-up approach, in which technologies process a large amount of data in order to continuously learn and improve their performance.[7] The analysis developed in this paper applies to both categories.

b) The IBA Working Group considers it risky to categorise different AI applications as 'green' or 'red' applications. With few exceptions (eg, autonomous weapons), most uses of AI cannot be identified as inherently bad or good for human rights. On the contrary, in most circumstances, AI can achieve positive social and economic objectives and there is a well-established link between digital technologies (such as AI) and the achievement of sustainable development goals (SDGs).[8] At the same time, algorithmic systems may have a negative impact on human rights, the rule of law and democracy, depending on the ways in which they are designed and deployed.[9] Generally speaking, 'red' AI could be characterised as AI where human rights risks are not properly identified and managed, and 'green' AI as AI where this has been done (for further information on the ways in which these risks can be addressed, please, see 5 and 7 below). Even though adverse impacts on human rights, rule of law and democracy may be intertwined,[10] this paper will focus exclusively on the risks of AI for human rights.

c) In the first instance, especially when AI systems are based on big data (ie, machine learning), obvious concerns arise regarding privacy and data protection.[11] AI systems are based on machine learning and their algorithms are only as good as their data. In order to operate at their prime, they require data, mainly 'Big Data'. The gathering, storage and usage of this data raise a range of human rights concerns. The concerns become even more acute when dealing with sensitive data, such as healthcare and biometric data (eg, facial recognition). This personally identifiable information is created daily by each of us (our digital footprint) and every time it is processed further personal data is created.

d) With regard to data gathering, the issues centre around consent, transparency and control over personal data (legibility, portability, negotiability, agency and democratisation of data). Until recently, wholesale gathering, storage, use and sale of personal data continued with very little user knowledge of the value and risks such big data held. Growing public concern and calls for action and answers, has led to closer inspection of the systems in place.[12] National authorities,

---

6 F Raso et al, *Artificial Intelligence and Human Rights: Opportunities and Risks*, Berkman Klein Center for Internet and Society Research Publication, 2018, available at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:38021439, last accessed 1 September 2020, p 10.

7 B G Buchanan, 'Can Machine Learning Offer Anything to Expert Systems?', *Machine Learning 4*, 1 December 1989, available at: https://doi.org/10.1023/A:1022646520981, last accessed 1 September 2020.

8 UN Human Rights Business and Human Rights in Technology Project (B-Tech), *Applying the UN Guiding Principles on Business and Human Rights to digital technologies*, November 2019. Available at: www.ohchr.org/Documents/Issues/Business/B-Tech/B_Tech_Project_revised_scoping_final.pdf; last accessed 28 August 2020.

9 F Raso et al, *Artificial Intelligence and Human Rights: Opportunities and Risks*, Berkman Klein Center for Internet and Society Research Publication. *CAHAI (2020) 06-fin.*

10 *CAHAI (2020) 06-fin.*

11 *CAHAI (2020) 06-fin*; J E Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 2000.

12 This brief report elaborates on some of the concerns and questions that need to be answered when it comes to algorithmic decisions, AI and personal data: C F Kerry, *Protecting privacy in an AI-driven world*, 10 February 2020, available at: www.brookings.edu/research/protecting-privacy-in-an-ai-driven-world, last accessed 7 September 2020.

regional bodies, international organisations, industry and civil society have embarked on a number of initiatives for better regulation of personal data use.[13] When it comes to personal data, privacy and security has become a central concern for individuals, companies and government authorities. Privacy forms part of the right to private life and is fundamental to the enjoyment of other human rights.[14] The gathering, storage and use of personal data interferes with privacy rights and measures must be put in place and closely monitored to ensure that privacy rights are not violated. Data protection laws exist in most countries recognising the right to respect for privacy.

e) Anonymised data sets are one way of limiting the privacy risks; however, the value of data often comes with the ability to identify or target an individual, as well as to draw correlations between different data sets and the individual.[15] Re-identification is also a possible concern. The richer the data, the more useful it is, but the greater the likelihood that an individual could be identifiable. Thus, anonymisation does not erase the need to conduct privacy risk assessments. This is also true of pseudonymised data, which may lower the risks, but also lower the value of the data.[16] This is why obtaining consent from the data subject is important for those who wish to gather and process personal data. The reason for collecting the data and its use must be made clear to users. If the use of the data changes, the company will need to update their privacy policy and get users to accept these new terms. However, most individuals do not read the privacy policy or statement before clicking accept. This is why privacy policies need to be monitored to ensure compliance with mandatory norms of data protection.

f) One way for this to occur is through law and regulation. The entry into force of the European Union's General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) has proven a useful tool to align data protection policies with the online transition of our lives and work. Without going into the details of the GDPR's operation, it is important to highlight that its approach is one which could be integrated into human rights and privacy due diligence assessments. The two step approach of setting out broad but clear principles ('lawfulness, fairness and transparency', 'purpose limitation', 'data minimisation', 'accuracy', 'storage limitation' and 'integrity and confidentiality'), coupled with the risk-based decision making (supported by the accountability requirement) is one which provides protection, whilst allowing flexibility so that responses can be tailored to a specific use and circumstance.[17]

---

13 See, for example: C Huet, *European Commission's Initiative in Artificial Intelligence*, available at: www.oecd.org/going-digital/ai-intelligent-machines-smart-policies/conference-agenda/ai-intelligent-machines-smart-policies-huet.pdf, last accessed 7 September 2020; European Commission's appointment of the High-Level Expert Group on AI, available at: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence, last accessed 7 September 2020; China's adopted strategy on AI in 2017, available at: https://flia.org/wp-content/uploads/2017/07/A-New-Generation-of-Artificial-Intelligence-Development-Plan-1.pdf, last accessed 7 September 2020; US Executive Order 13859 in 2019, available at: www.hsdl.org/?abstract&did=821398, last accessed 7 September 2020.

14 As guaranteed under Article 8 of the European Convention on Human Rights and other international and regional human rights treaties. For further analysis of the human rights implication, see the *Council of Europe Study on the Human Rights Dimensions of Automated Data Processing Techniques* (in particular algorithms) and *Possible Regulatory Implications*, DGI (2017)12, available at: https://rm.coe.int/algorithms-and-human-rights-study-on-the-human-rights-dimension-of-aut/1680796d10, last accessed 7 September 2020. See also M Scheltema, *Embedding private standards in AI and mitigating artificial intelligence risks*, in Proceedings – 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People and Smart City Innovation, SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019 (pp 305–310), digital object identifier (DOI): 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00096.

15 A simple example is where you use a search engine to find a product: that piece of information is gathered and used by other platforms and sites trying to sell you a similar product; you may also receive an email with suggestions.

16 Article 4 (5) GDPR defines pseudonymisation as meaning 'the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.'

17 For a summary of the role of the GDPR in regulating the use of Big Data see: T Christou and I Walden, *Legal and Regulatory Implications of Disruptive Technologies in Emerging Market Economies*, 28 June 2018, available at: https://ssrn.com/abstract=3230674; or http://dx.doi.org/10.2139/ssrn.3230674, last accessed 7 September 2020.

g) Another way is through self-regulation, including conducting human rights due diligence as recommended by the UNGPs. The recent stance of Apple and Google over the track and trace apps, which governments wanted to introduce as a response to the coronavirus pandemic, may have restored some faith in the ability of companies to protect our personal data. Their insistence on a decentralised database limited the exposure of user personal data.[18] The reality is that, in a globalised world, it is necessary for public and private actors to work together, to develop standards, practices and checks to protect personal data. The technology and AI not only create risks which must be mitigated but also offer the solutions, whether through privacy by design or the ability to speedily, accurately and autonomously identify and either rectify or delete personal data.

h) AI systems are vulnerable to cyberattacks that can compromise the processing of incoming data and data sets required to instruct the AI for generating a specific output. This can distort the specific output by responding incorrectly to the data inputs and risk the human rights due diligence process. It is essential to have a cybersecurity due diligence process that incorporates the review of the governance, processes and controls for securing the information assets used for the development of AI systems. Cybersecurity due diligence is based on risk management programs, compliance obligations and best practices and needs to be incorporated at the beginning of the design phase of AI system development to avoid causing any harm.

i) Additional human rights concerns may be associated with the way in which algorithms are designed and with the type of data used to train them. For example, facial recognition may facilitate secure access to mobile phones, a house or an office, but may also have negative human rights consequences if used in law enforcement settings. In particular, this specific function has raised concerns in relation to the potential risks of discrimination. For reasons of commercial sensitivity, the data used to train facial recognition software is rarely shared with public authorities. This lack of transparency makes it difficult for end users to verify whether the software contains inherent biases against a specific ethnic group. For this reason, the Court of Appeal of England and Wales has concluded, in a recent judgment, that there is currently no adequate legal framework for the use of facial recognition technology to support police activities.[19] Similarly, credit-scoring algorithms often deployed to assess the credit worthiness of individuals may perpetuate and amplify patterns of inequality.[20] The risks to human rights themselves may also vary depending on the autonomy of AI. If AI has an advisory function, these risks may be more limited, compared with autonomous AI in which no humans are involved (eg, self-driving vehicles).

j) Finally, in all circumstances, AI systems do not exist in a vacuum and their impact on human rights may depend on the social economic contexts in which they operate, as well as on the nature of their end users. This is particularly evident with private surveillance products, whose successful deployment is generally operated without the knowledge and consent of end users.[21]

---

18 M P Sacco et al, *Digital contact tracing for the Covid-19 epidemic: a business and human rights perspective*, June 2020, IBA Report, available at: www.ibanet.org/Article/NewDetail.aspx?ArticleUid=4b11819d-c580-47fe-b680-19bdbc201328, last accessed 6 September 2020.

19 *R (Bridges) v Chief Constable of South Wales* [2020] EWCA Civ 1058.

20 F Raso et al, *Artificial Intelligence and Human Rights: Opportunities and Risks*, Berkman Klein Center for Internet and Society Research Publication, 2018, available at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:38021439, last accessed 1 September 2020, p 28.

21 UN Human Rights Council, *Surveillance and human rights: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 28 May 2019, p 7.

In addition to the risks for the right to privacy, these instruments may be adopted by repressive governments in an effort to silence dissent, with a clear impact on freedom of association and freedom of expression.[22]

'*5 – Mapping of instruments applicable to artificial intelligence (v. International legal instruments, ethical guidelines and private actors)*' *and* '*7 – Main elements of a legal framework for the design, development and application of artificial intelligence (i. Role and responsibilities of member states and of private actors in developing applications which are in line with such requirements; ii. Liability for damage caused by artificial intelligence)*'

k) Considering the potentially disruptive impact of AI on human kind,[23] the current debate is focusing on the governance structure better suited to promoting innovation, while also reducing the negative risks related to these technologies.[24] Most initiatives have adopted an ethics-based approach to achieve this balance,[25] as, among others, the one led by the European Commission High-Level Expert Group on AI.[26] However, these ethical frameworks do not always make a clear distinction between more legally binding 'ethical' obligations, for example those emerging from human rights and prescribed by law, and ethical considerations which do not have such a basis. These two types of ethical issues should be differentiated and it may be advisable to reserve the term 'ethical guidelines' to those instruments which do not have a legal basis as yet.

l) An important shift towards a regulatory/human rights approach to AI is represented by the recently published *Recommendation of the Council of Europe on the human rights impacts of algorithmic systems.*[27] This document provides a set of guidelines for governments and private actors to consider when designing, developing and deploying AI systems and refers to the '*Protect, Respect and Remedy'* framework set out by the UNGPs. Together with the Organisation for Economic Cooperation and Development (OECD) *Guidelines for Multinational Enterprises* ('OECD Guidelines')[28] and the OECD *Due Diligence Guidelines for Responsible Business Conduct*[29] ('OECD Due Diligence Guidelines'), the UNGPs represent the most authoritative global standards to address business impact on all human rights, applicable to both states and businesses. Although these are general principles and are not specifically targeted at AI, the risks management and remedial processes set out in these documents and designed to be applied by all business sectors (human rights due diligence) are relevant to AI as well.

m) As far as the role of the private sector is concerned, particular attention should be paid to the role of human rights due diligence. According to Principle 15 of the UNGPs, through this process, business entities are expected 'to identify, prevent, mitigate and account for how they

22 *Ibid*, p 8.

23 O Osoba and W Welser IV, *An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence*, (2017), Rand Corporation available at: www.rand.org/content/dam/rand/pubs/research_reports/RR1700/RR1744/RAND_RR1744.pdf, last accessed 7 September 2020.

24 J Cowls, L Floridi, *Prolegomena to a White Paper on an Ethical Framework for a Good AI Society*, (2018), Oxford University, available at: http://dx.doi.org/10.2139/ssrn.3198732, last accessed 7 September 2020.

25 *Toronto Declaration on Protecting the Rights to Equality and Non-Discrimination in Machine Learning Systems*, 2018, available at: www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf, last accessed 7 September 2020.

26 See, for instance, European Commission High Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI High-Level Expert Group on Artificial Intelligence*, (2019), available at: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence, last accessed 3 September 2020.

27 Council of Europe, *Draft Recommendation on the human rights impacts of algorithmic systems*, (2019), available at: https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154, last accessed 3 September 2020.

28 OECD, *OECD Guidelines for Multinational Enterprises*, 2011, OECD Publishing, available at: www.oecd.org/daf/inv/mne/48004323.pdf, last accessed 3 September 2020.

29 OECD, *OECD Due Diligence Guidance for Responsible Business Conduct*, 2018, available at: http://mneguidelines.oecd.org/OECD-Due-Diligence-Guidance-for-Responsible-Business-Conduct.pdf, last accessed 3 September 2020.

address their impacts on human rights'.[30] This activity should focus on the risks to human rights, rather than to business activities and should be conducted on an ongoing basis, since 'the human rights risks may change over time as the business enterprise's operations and operating context evolve'.[31]

n) The revised 2011 OECD Guidelines have expanded the scope of due diligence beyond human rights to incorporate broad risks (including bribery and taxation) to responsible business conduct (RBC). RBC due diligence, as defined by the OECD Due Diligence Guidelines, includes six steps as displayed in the following diagram (Figure 1):[32]

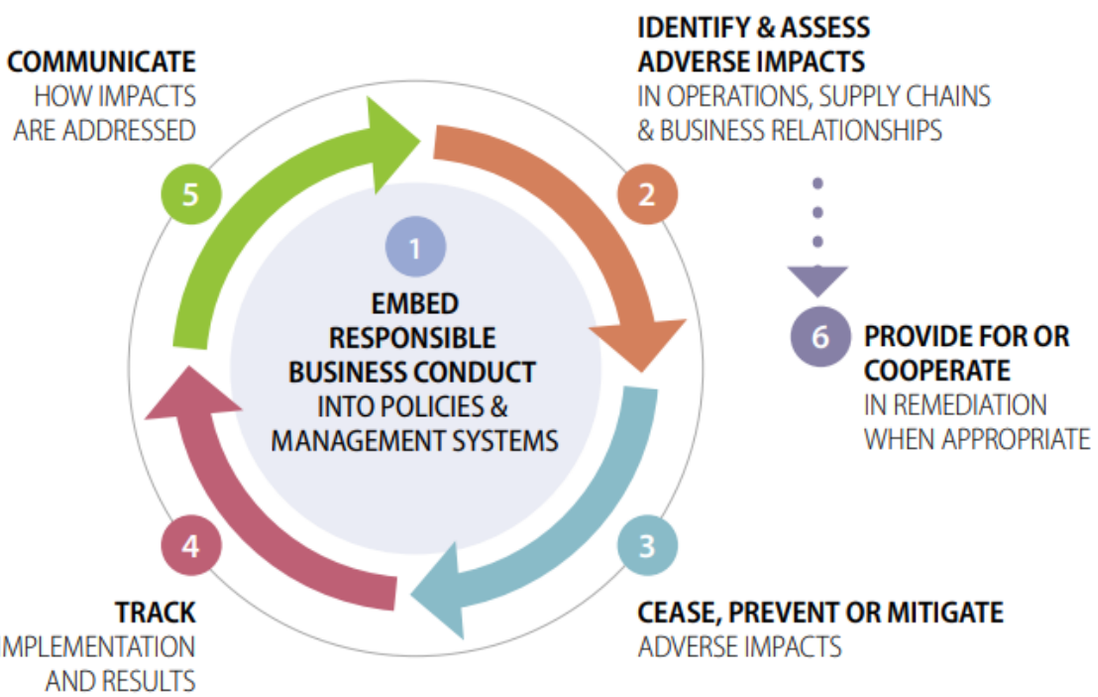

Figure 1

An internationally operating company should draw up an international RBC policy and integrate this into all its policy and management systems. On the basis of this RBC policy, an ongoing due diligence investigation needs to be carried out, which primarily consists of identifying and analysing risks related to RBC themes. After identifying and analysing these risks and determining the company's level of involvement, it needs to take appropriate measures to prevent or limit the risks from materialising. To improve the due diligence process and the measures to be taken, the company needs to monitor the effectiveness of its due diligence policy and communicate the progress made with its stakeholders. If the company has caused or contributed to adverse effects, it should provide for an appropriate remedy. This can be achieved through legitimate processes, including judicial and non-judicial complaint mechanisms (see section (p) below).

30 UNGPs, Principle 15 (b).

31 *Ibid*, Principle 17 (c).

32 Figure 1: Due Diligence Process and Supporting Measures in OECD Due Diligence Guidance for Responsible Business Conduct, p 21, available at: http://mneguidelines.oecd.org/OECD-Due-Diligence-Guidance-for-Responsible-Business-Conduct.pdf, last accessed 3 September 2020.

o) In 2019, the UN Office of the United Nations High Commissioner for Human Rights (OHCHR) launched a project to explore how the UNGPs may be relevant to the tech sector ('B-tech project').[33] From its scoping paper, the important role attributed to human rights due diligence appears evident.[34] Moreover, in the context of AI, the risk identification and management not only applies to developers of AI but also to its users, regardless whether they are private or public users. In particular, the following activities have been regarded as presenting higher risk of having adverse impacts on human rights:[35]

- gathering of large volumes of data (either to train algorithms or to sell insights to third parties);
- selling products to, or partnering with, governments seeking to use new technologies for state functions or public service delivery that could disproportionately put vulnerable populations at risks;
- the promise of hyper-personalisation in human resources or marketing decisions, which could lead to discrimination;
- using 'algorithmic bosses' to mediate the relationship between workers and firms that generate business value from the offline work being done, while limiting labour protections for those workers; and
- models that are informed by, or inform, the personal choices and behaviours of populations without their knowledge and consent.

According to the UNGPs, human rights due diligence should look at the 'adverse human rights impacts that the business enterprise may cause or contribute to through its own activities, or which may be directly linked to its operations, products or services by its business relationships'.[36] This implies that human rights due diligence goes beyond a company's products and services, to include all the value chain. According to the scoping paper of the B-tech project, software developers and producers of digital technologies (including AI) have to look, in particular at end users, 'as it is mostly in their use that human rights harms will manifest'.[37] In this context, companies are also expected to exercise leverage with their business partners and users, including through multi-stakeholder initiatives.[38]

p) Both the UNGPs and the OECD Guidelines provide for judicial and non-judicial remedies to address adverse human rights impacts associated with business activities.[39] In particular, the UNGPs identify the following three categories: state-based judicial mechanisms; state-based non-judicial mechanisms and non-state-based non-judicial mechanisms.[40] These principles highlight

---

33 Office of the United Nations High Commissioner for Human Rights, *UN Human Rights Business and Human Rights in Technology Project (B-Tech Project)*, further information available at: www.ohchr.org/EN/Issues/Business/Pages/B-TechProject.aspx, last accessed 3 September 2020.

34 Office of the United Nations High Commissioner for Human Rights, *UN Human Rights Business and Human Rights in Technology Project (B-Tech), Applying the UN Guiding Principles on Business and Human Rights to digital technologies, Overview and Scope*, November 2019 ('UN B-Tech Scoping Paper'), p 5, available at: www.ohchr.org/Documents/Issues/Business/B-Tech/B_Tech_Project_revised_scoping_final.pdf, last accessed 3 September 2020.

35 *Ibid* n 34, p 7.

36 UNGPs, Principle 17.

37 UN B-tech scoping paper, p 6.

38 *Ibid.*

39 UN B-Tech Scoping Paper, p 8.

40 UNGPs, Principles 26–31.

that state-based mechanisms are often insufficient and they should represent the 'foundation of a wider system of remedy'.[41] This is particularly true in the context of digital technologies (including AI), where abuses may be the result of decisions made by algorithms; affected rights-holders may represent hundreds of millions and the value chain may be composed of dozens of companies, rather than a single actor.[42] For example, the Cambridge Analytica scandal has shown that, while many people can be affected by improper provision of data to third parties (which may have considerable impact on democratic values and the rule of law), the individual damage is too limited to justify legal action and governments do not have the necessary tools to address these challenges.

q) Notwithstanding the important role played by the UNGPs and the OECD Due Diligence Guidelines, criticisms have been raised by stakeholders that their voluntary nature may lead, in some circumstances, to a 'box-ticking' exercise.[43] In the attempt to give teeth to these non-binding instruments, some countries have enacted laws to mandate due diligence for general application (eg, France)[44] or more specific issues (eg, the EU regulation on conflict minerals; UK Modern Slavery Act).[45, 46] In addition, the EU is currently considering introducing mandatory human rights and environmental due diligence in all sectors[47] by 2021.

r) Human rights due diligence is not a one-size-fits all exercise and the specific features of the market in which business entities operate have to be taken into account when carrying out this activity. The OECD has published sector-specific guidelines on RBC due diligence to support the work of the private sector in this respect.[48] In addition, some multi-stakeholder initiatives, some of which are legally binding agreements (such as the Dutch International Responsible Business Agreements)[49] involve companies, trade unions and governments (in the case of the Dutch Agreement) in which signatory companies commit themselves to identifying and addressing human rights risks in their supply chain. In this context, arbitration can be used to address business-to-business disputes, as well as business-to-rights-holders disputes (or their representatives, like trade unions).

s) When looking at digital technologies and, in particular, at AI applications, practical standards on the way in which human rights due diligence can be implemented in practice are currently missing. **The IBA, together with lawyers from business, academia and law firms, has established a project to identify and elaborate on ways in which human rights due diligence can be implemented through legal instruments in practice in the AI area (for example in contracts) and how dispute resolution should be shaped in connection with it**. This project will include holding roundtables with relevant stakeholders (including businesses) and engaging with international

41 UN B-Tech Scoping Paper, p 8.

42 *Ibid* n 41.

43 European Commission, *Study on due diligence requirements through the supply chain*, 2020, p 111, available at: https://op.europa.eu/en/publication-detail/-/publication/8ba0a8fd-4c83-11ea-b8b7-01aa75ed71a1/language-en, last accessed 4 September 2020.

44 This is the French Duty of Vigilance law, available at: www.senat.fr/leg/ppl14-376.html, last accessed 7 September 2020.

45 Regulation EU 2017/821, available at: https://publications.europa.eu/en/publication-detail/-/publication/8b0e378b-3c59-11e7-a08e-01aa75ed71a1/language-en/format-PDFA1A, last accessed 7 September 2020.

46 UK Modern Slavery Act, available at: https://www.legislation.gov.uk/ukpga/2015/30/contents/enacted, last accessed 7 September 2020.

47 See, for example, R McCorquodale and M Scheltema, *Core elements of EU mandatory human rights and environmental due diligence*, available at: www.business-humanrights.org/en/expert-contribution-core-elements-of-an-eu-regulation-on-mandatory-human-rights-environmental-due-diligence, last accessed 4 September 2020.

48 *OECD Guidelines for Multinational Enterprises*, available at: http://mneguidelines.oecd.org/sectors, last accessed 7 September 2020.

49 Agreement on International Responsible Business Conduct, available at: www.imvoconvenanten.nl/en, last accessed 7 September 2020.

organisations, such as the OECD, the UN (in particular, in relation to the B-tech project) and CAHAI. These outcomes may be relevant for CAHAI and the IBA is willing to collaborate with CAHAI on this, for example by inviting its representatives to these roundtables.

'*9 – Possible practical mechanisms to ensure compliance and effectiveness of the legal framework (such as for instance the creation of mechanism of ex-ante verification and/or certification, oversight by independent authorities, sandboxing, etc)*'

t) Corporate compliance with UNGPs, OECD Due Diligence Guidelines and private sector standards can be verified by third-party certification bodies. In a globalised market economy with complex value chains, these entities may play an important role in reducing information asymmetries between companies and stakeholders. For these reasons, since the 1990s these schemes have proliferated in diverse sectors, from health and safety to social and anti-bribery standards.[50] A similar instrument, focusing on AI applications, is currently being developed by the International Organization for Standardization (ISO) and should be published by the end of 2021.[51] In particular, ISO SC42 standard aims to define a risk-management system suited to identify, reduce and address the risks associated with AI applications, including human rights.[52] Beyond this, CEN/Cenelec are also exploring whether a European standard, with even more elaborate ethical/human rights principles should be developed.[53]

u) However, certification schemes are not immune from errors and mistakes, in particular, when it comes to the assessment of compliance with social standards.[54] In many countries, severe competition exists between certification bodies and the lack of regulation on the certification process in this field may lead to a race to the bottom in the provision of social auditing services.[55] This situation is exacerbated due to the fact that certification bodies are often remunerated by the same entities subject to verification, with clear consequences for the independence of this process.[56] Furthermore, certification is costly and procedural aspects (such as, the review of documentation) might prevail over assessing actual improvement of quality in an organisation or changes in environmental impact.

v) To conclude, certification schemes can support the regulatory framework promoting an active private sector in the prevention and management of the risks of adverse human rights impacts associated with AI applications, if the following conditions are satisfied: certifications should not be used by companies to shield themselves from potential liability claims, associated with their conduct; and the certification process is subject to regulation regarding the qualification of auditors; the standards adopted and the ways in which conflicts of interest are managed. Preferably, outcomes of non-state based grievance mechanisms should be aggregated (not revealing individual complaints or complainants but, for example, building on the nature and

50 R Van Tulder et al, *From chain liability to chain responsibility*, Journal of Business Ethics 2009, p 399.

51 International Standardization organization (ISO), ISO/IEC JTC 1/SC 42 *Artificial intelligence*, available at: https://www.iso.org/committee/6794475.html, last accessed 4 September 2020.

52 *Ibid.*

53 CEN and CENELEC launched a new *Focus Group on Artificial Intelligence*, 16 May 2019, available at: https://www.cencenelec.eu/news/articles/Pages/AR-2019-001.aspx, last accessed 4 September 2020.

54 See, for example, M Scheltema, *Assessing effectiveness of international private regulation in the CSR arena*, Richmond Journal of Global Law and Business 2014, vol. 13, pp 324 and 325.

55 T Van Ho and C Terwindt, *Assessing the Duty of Care for Social Auditors*, European Review of Private Law, 2019.

56 J Coffee, *Gatekeepers: The Professions and Corporate Governance*, Clarendon Lectures in Management Studies, 2006.

number of complaints) and fed into the certification process; and standards and certification should strive for continuous improvement (ie, the standards should incentivise such improvement, for example, by being adapted to technological changes and through regular evaluation with all relevant stakeholders).

Conclusions

The IBA Working Group would like to thank CAHAI for this opportunity to comment on the Draft Feasibility Study and would like to submit the following recommendations:

- The UNGPs, the OECD Guidelines and the OECD Due Diligence Guidelines should be included in sections 5 and 7 of the Draft Feasibility Study. In particular:

  (i) when defining the role and responsibility of private actors (Section 7, ii), the Draft Feasibility Study should refer to the responsibility of business entities to carry out human rights due diligence, as defined by the UNGPs and the OECD Due Diligence Guidelines; and

  (ii) when looking at 'liability for damage caused by AI' (Section 7, iii), the Draft Feasibility Study should go beyond state-based judicial remedies and should refer to state-based non-judicial remedies, as well as to non-state-based non-judicial remedies.

- In order to avoid excessive regulatory burdens on the private sector, the Draft Feasibility Study should look at the existing provisions mandating human rights due diligence and, in particular, at the upcoming EU Legislation on mandatory human rights and environmental due diligence.

  In order to promote a more effective implementation of human rights due diligence, the specific challenges of AI should be considered and private actors, as well as relevant stakeholders, should be consulted. **The IBA Working Group is working on the publication of a set of due diligence standards on AI and is keen to engage with CAHAI in this regard.**