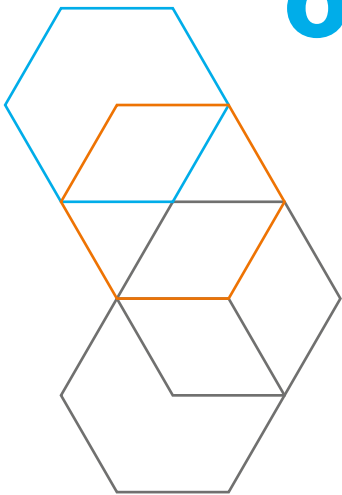# Mapping
the challenges
and opportunities
of artificial intelligence
for the conduct
of diplomacy

DiploFoundation

DiPLO
AILab

DiPLO
www.diplomacy.edu

Table of Contents

Acknowledgements

Executive summary

This report provides an overview of the evolution of diplomacy in the context of artificial intelligence (AI). AI has emerged as a very hot topic on the international agenda impacting numerous aspects of our political, social, and economic lives. It is clear that AI will remain a permanent feature of international debates and will continue to shape societies and international relations. It is impossible to ignore the challenges – and opportunities – AI is bringing to the diplomatic realm. Its relevance as a topic for diplomats and others working in international relations will only increase.

Understanding AI and mapping the relationship between AI and diplomacy

Despite the recent hype, AI is not a new phenomenon. There is also no universally agreed definition of AI and it acts as an *umbrella term* for machine learning, automated reasoning, robotics, computer vision, and natural language processing (NLP). AI is also a *moving target*; the more integrated it becomes in everyday life, the less likely we are to still refer to it as AI.

Discussions about AI in the context of foreign policy and diplomacy often lack clarity in terminology. However, an understanding of the opportunities and limits of the current state of AI, machine learning in particular, is important in order to build realistic expectations and navigate opportunities as well as challenges.

When trying to understand the relationship between AI and diplomacy, it is useful to build on the distinction between AI as a diplomatic topic, AI as a diplomatic tool, and AI as a factor that shapes the environment in which diplomacy is practised.

- As a *topic for diplomacy*, AI has become a relevant concern for many debates. It is relevant for topics ranging from economy and business, and security, all the way to democracy, human rights, and ethics.
- AI as a *tool for diplomacy* looks at how AI can support the functions of diplomacy and the day-to-day tasks of diplomats. It is important to remember that AI is not an end in itself, but is only as useful as it can meaningfully support diplomats.

- As *a factor that impacts the environment in which diplomacy is practised*, AI could well turn out to be the defining technology of our time and as such determine economic, social, and political successes. Depending on a country's ability to harness the technology, military and economic balances could shift. In military terms, Lethal Autonomous Weapons Systems (LAWS) pose a particular concern.

The relationship between AI and diplomacy can also be framed as in the case of science diplomacy from which a three-part typology emerges:

- First, *diplomacy for AI* serves as a reminder that diplomacy has a role to play in supporting AI research by fostering co-operation across borders and supporting existing institutions.
- Second, *AI for diplomacy* alludes to the possibility that international scientific co-operation in the field of AI can contribute to improved relations between countries. The co-operation between diplomats can build bridges when official relations between countries are strained.
- Third, *AI in diplomacy* parallels the concept of AI as a tool for diplomacy. Additional insights generated through AI applications can contribute to better foreign policy decision-making.

National AI strategies and international co-operation

Over the last two years, since 2017, a number of countries have begun to prepare or release national AI strategies. Developing a national AI strategy is a must in order to be prepared for the impact AI will have on international relations.

Many countries emphasise the importance of international co-operation on AI in their national strategies and reports. Their main reasons for doing so include the need for pooling resources for research and development, the need for access to (big) data, a desire to set standards and develop an ethical framework, and a desire to build on and amplify shared norms and ethics.

The European Union (EU) focuses on closer co-operation and co-ordination in a variety of forms. It also aims to set global quality standards for AI applications and to become more active in multilateral discussions, especially where the use of AI for military purposes is concerned.

The UN is concerned about a widening gap between developed and developing countries when it comes to the ability of the latter to harness opportunities related to AI. This is particularly important because of the potential of AI to support the implementation of the sustainable development goals (SDGs). Those countries taking the lead in AI could consider including capacity building in the area of AI as part of their development co-operation.

AI as a tool for diplomacy

To think through AI's potential to serve as a tool for diplomats, the distinction between assisted, augmented, and automated intelligence is a useful starting point. At this point in the development of AI, it is best used to assist and augment intelligence.

Since a substantial part of diplomatic practice is concerned with textual data, for example in the form of treaties and diplomatic reports, tools that can support the meaningful analysis of this data at scale, in particular AI that uses NLP techniques, are of particular interest.

The analysis of texts at scale has the potential to make the work of diplomats more effective and free up time and resources. As a consequence, more time can be spent on aspects of diplomatic work that require uniquely human skills and human intuition.

Ministries of foreign affairs (MFAs) need to be prepared to engage and build partnerships with software companies to obtain these tools and to adapt them to their needs. Customised solutions build on close co-operation and it is, therefore, important to think about the availability of personnel to engage with technological companies.

Addressing the human rights dimension of AI

Looking at one particular topic on the diplomatic agenda that will be impacted by AI applications, we zoom in on the human rights dimension of AI.

When we talk about the human rights implications of AI, we are in a lot of ways implicitly referring to the unprecedented need for data when it comes to building AI. Generally speaking, data in all its forms (big data, open data, personal data, sensitive data) is a critical juncture in understanding human rights with regard to AI.

Irrespective of where it may be found, it is the *potential* of AI to discriminate and infringe on human rights which matters, even if there is no underlying intention to do so. The rights to privacy, family, home, and correspondence; the prohibition of discrimination; and the right to hold opinions and to freedom of expression are areas that will be particularly impacted by AI.

It is the duty of states to encourage the private sector to design, develop, and deploy AI with respect for human rights. States are the primary guarantors of these rights and freedoms and are formally obligated to protect them.

On this basis, any state claiming a pioneer role in the development and/or reliance on new technologies bears a special responsibility for striking the right balance between rights, duties, and interests.

Next steps for diplomacy and MFAs

There are three next steps that MFAs can take in order to get ready to address the challenges and opportunities of AI effectively.
- First, it is useful to develop indicators that help track the efforts made by MFAs in advancing national AI strategies. When it comes to the MFA's engagement with AI, this is important in order to allocate resources, track progress, document lessons learned, and stay accountable to citizens.
- Second, there is a need to engage in capacity building that includes not just those working directly with topics related to AI. Every diplomat needs to have a basic level of understanding of the technology.
- Third, organisational adjustments include (a) establishing an agile, cross-cutting unit tasked with exploring the opportunities of using AI as a tool for diplomacy and (b) identifying AI champions across ministerial departments who can further the dialogue on various aspects of AI within the ministry.

Introduction

Artificial Intelligence (AI) has, without doubt, entered national and global discussions. We hear almost daily about new advances in AI and new uses of the technology. We also hear about its dangers and receive warnings that it will soon pose an existential threat to humanity. It is difficult to find a way to navigate between the hype and the dystopian vision. It is clear, however, that AI applications are already all around us. AI is here to stay and its role in our everyday lives will only increase.

Discussions about the opportunities and challenges of AI have entered the various sectors of society, the economy, and politics. People see clear potential related to, for example, economic growth and more tailored and efficient public services, but they also see dangers related to, among other things, discrimination and the loss of jobs.

AI is often talked about as a technology with great potential for disruption and diplomacy will have to adapt. But it is safe to say that diplomacy is here to stay. In fact, looking at the fundamental questions and dilemmas that AI and its applications pose, it is clear that we need diplomacy more than ever. Diplomats are needed to foster understanding between countries with very different approaches to AI and to build and maintain relationships across borders with regard to, for example, shared understandings, the peaceful use of the technology, scientific co-operation, and ethical questions.

This report follows on from a previous DiploFoundation study on *Data Diplomacy: Updating diplomacy to the big data era*. To transition from researching the relevance of big data for diplomacy to looking at the interplay between AI and diplomacy was a natural progression for us. In doing so, we move our attention from a focus on the data needed to inform decision-making to the processes and tools involved in making sense of that data and arriving at recommendations for decisions.

With this report, we present an inception study on AI and diplomacy that intends to map the field. While debates on AI and diplomacy and foreign policy are slowly emerging, we feel that it is important to provide a broad overview on the basis of which further, more detailed debates can take place and further research into more specific questions can emerge. In this sense, the topics covered in this report are meant as food for thought and as suggestions for further investigation. At the same time, we also give concrete advice and make suggestions for how diplomats and ministries of foreign affairs (MFAs) can prepare to deal effectively with the challenges and opportunities posed by AI.

In this report, we explore the relationship between AI and diplomacy in four main steps. The *first chapter* provides an overview by asking two key questions: a) How can we make sense of AI? and b) How can we map the relationship between AI and diplomacy? By providing both a simple understanding of AI and by suggesting two typologies for mapping the relationship between AI and diplomacy, this chapter lays the groundwork for subsequent ones.

The *second chapter* provides an overview of national AI strategies and related reports that have been published over the last two years. As countries try to position themselves internationally, the differences and similarities in their approach to AI become visible. In particular, this chapter looks at the way in which (big) data, the key ingredient for machine learning AI, is approached in the strategies and documents. It also looks at aspects of competition and co-operation between countries and highlights reasons for co-operation.

The *third chapter* shifts the focus from how AI is discussed to how AI is used in the context of diplomacy by looking at AI as a tool for diplomatic practice. This chapter starts by outlining different ways in which AI, understood as a tool for diplomacy, can be approached. It describes one case study from the field of natural language processing (NLP) in greater detail. Utilising AI as a tool for diplomacy also necessitates a closer engagement with software companies and the building of partnerships between MFAs and these companies. Hence, in the second half of this chapter, we give advice, based on our own experience, on how to approach these partnerships.

We then move from AI as a tool for diplomacy to AI as a topic in diplomatic practice. AI impacts numerous areas in the social, economic, and political fields. In the *fourth chapter*, we zoom in on human rights and AI. The discussion looks at the rights to privacy, family, home, and correspondence; the prohibition of discrimination; and the right to hold opinions and to freedom of expression.

It explains in what ways these are particularly impacted by AI.

By way of *conclusion*, we add four additional recommendations for MFAs and diplomats who want to get ready to effectively deal with the impact of AI on international agendas and diplomatic practice.

1. Making sense of AI and mapping the relationship with diplomacy

Sometimes, we may forget how much we owe to flakes and wheels, to sparks and ploughs and to engines and satellites [but then][w]e are reminded [again] of such deep technological debt.[1]

The main lesson of thirty-five years of AI research is that the hard problems are easy and the easy problems are hard. The mental abilities of a four-year-old that we take for granted—recognising a face, lifting a pencil, walking across a room, answering a question—in fact solve some of the hardest engineering problems ever conceived.[2]

Finding a way to navigate between hype and dystopian vision when it comes to artificial intelligence (AI) and beginning a nuanced discussion of the potential and impact of AI for diplomacy necessitates a clear understanding of the technology. To develop such an understanding, this chapter serves a dual purpose: first, to offer a broad-based understanding of AI in order to make sense of it from a non-technical perspective and second, to think through the relationship between AI and diplomacy in a way that provides a useful framing for the conceptual work and planning of practitioners.

There is a growing body of literature on AI and its impact on the political, economic, and social spheres. However, there is also a tendency to use broad generalisations without unpacking key definitions or distinguishing between various forms of AI. Working towards a more nuanced understanding is crucial if we are to navigate between utopian visions of a bright future and their dystopian counterparts. It will be important to find a practical path for diplomats and foreign policy practitioners that recognises AI in the various ways in which it relates to the practice of diplomacy.

Over the last two years in particular, since 2017, we have seen a number of governments issuing statements and reports on AI. There is also a nascent literature on the foreign policy implications of AI. International organisations have begun to look at the implications of AI for their work. Yet, a comprehensive framework for grasping the relationship between AI and diplomacy in its complexity and in its entirety is still missing.

Building on these observations, this chapter gives, first, a conceptual and practical overview of AI in order to lay a foundation for understanding on which the following parts can build. And, second, it suggests a mapping of the relationship between AI and diplomacy.

1.1 Making sense of AI

To fully grasp the relationship between AI and diplomacy, a clear understanding of what AI is and what it is not and what we can and cannot expect from the technology at this point in time, offers a crucial starting point. However, there are two challenges associated with providing such an overview.

First, AI is a moving target. It is often the case that technology that has become common place is no longer described as AI.[3] The term AI seems to be reserved for the cutting edge of current technological standards and for projections into the future, not for the kinds of things that computers have become routine at. Second,

AI is a 'suitcase' term. 'It's a concept that appears simple enough but is actually endlessly complex and packed – like a suitcase – with lots of other ideas, concepts, processes and problems.'[4] Speaking about AI in very broad terms without reference to specific examples, introduces vagueness into the discussion that can make it hard to have a meaningful conversation about the social, political, and economic impact of AI. At the same time, it is challenging for non-experts to make the necessary distinctions between different aspects of AI, especially with regard to the underlying algorithms and their functioning. Despite these challenges, if we want to move this discussion forward, we need to strive for as much conceptual clarity as possible.

Defining AI

In very broad terms, we can say that AI is 'the scientific study of the computational principles behind thought and intelligent behaviour'.[5] Other definitions use human intelligence and capabilities as a point of departure and describe AI as, for example, '[t]he theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages'.[6]

It is most useful to remind ourselves that AI is not a new phenomenon. The 1956 Dartmouth Summer Research Project on Artificial Intelligence is considered the birthplace of AI as we know it today. The proposal for the conferences contains the assumption that 'every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.'[7]

Since then, the field of AI as a scientific study has become a lot more diverse and has branched out into an increasing number of sub-fields. And although some of the wilder predictions about future developments in AI have not become a reality, important advancements have been made.

Over the last few years in particular, a number of highly publicised events showcased the power of AI in relation to human cognitive capabilities. In 2011, IBM Watson defeated human champions in the TV game show, *Jeopardy.*[8] In 2016, AlphaGo (a computer program that plays the boardgame *Go*) won against the *Go* world champion. Last year, IBM introduced Project Debater, an AI which debates humans on complex issues.

While these events highlight the steps taken in the advancement of AI, it is crucial to consider a number of fundamental distinctions when beginning to discuss the implications of AI for foreign policy and diplomacy. In the following sections, we look at the distinction between weak and strong AI, and generalised and specialised AI, and highlight machine learning as the key that led to recent advances in the technology. It is also important to recognise that AI is an umbrella term, encompassing a number of different approaches and techniques.

The philosopher John Searle developed the distinction between weak and strong AI. He asked the question whether computers can have a mind (strong AI) or whether they can merely simulate one (weak AI). He concluded that strong AI is not possible. Computers might appear intelligent by performing certain tasks, but they lack a deeper understanding of what it is that they are doing.[9] This is important to emphasise: While certain tasks are being performed by an AI system, there is no sense of context or implication. A system analysing satellite images of refugee camps and a system analysing the movement of combatants on the ground have, unsurprisingly, no sense of the very different implications these two cases have in a foreign policy context. Weak AI can be understood as 'the use of software to accomplish *specific problem solving or reasoning* tasks'.[10]

The distinction between generalised and specialised AI builds on a similar distinction. A specialised AI system is trained for a specific purpose with very little meaningful ability to accomplish other tasks. This means that trying to use the same model or system for another purpose is either not possible or leads to very questionable results. While it has been reported that AlphaGO is now also proficient at two other boardgames, we are very far from a generalised AI.[11] Vint Cerf, often referred to as one of the fathers of the Internet, expressed it in the following way.

> 'AI normally stands for artificial intelligence, but I've often concluded it stands for "artificial idiot." And the reason is very simple. It turns out that these systems are good, but they're good in kind of narrow ways. And we have to remember that so we don't mistakenly imbue some of these artificial intelligences and chat bots and the like with a breadth of knowledge that they don't actually have, and also with social intelligence that they don't have.'[12]

From this point of view, it is clear that current AI is no replacement for humans except in areas in which a high degree of standardisation, and thereby automation, is possible.

However, there are debates about how to move from specialised to generalised AI. Some propose to think of it as hierarchies of intelligence, in which current versions of AI serve as lower levels to be eventually built up to higher levels. At the same time, others argue that a completely different path is needed to take us from specialised to generalised AI. Within this debate, it is worth keeping in mind the fact that research into generalised AI exists is not the same as saying that generalised AI exists.

Areas of AI[13]
Automated reasoning is concerned with using programs for deduction (solving mathematical problems for example), planning, and optimising.
Robotics focuses on building machines that can interact with the world. Automating factory processes is a good example of the use of robotics with implications for the economic sector. Machines can take over jobs from human beings. 'Dark factories', factories where no light is required because humans do not work there, have emerged.
Computer vision is about making machines 'see'. This is, for example, highly relevant for autonomous vehicles who need to understand and navigate the environment around them and react quickly to changes.
Natural language processing (NLP) deals with understanding and using natural language, which means understanding the way humans speak and write and being able to do so as well. Applications include understanding and answering questions, translating between natural languages, and recognising speech.

While this might seem more like a debate for philosophers and computer scientists than for diplomats and policymakers, it is important to be aware of this distinction. It makes clear that all current applications of AI are forms of specialised AI. In this area, we can expect to see advancements to the extent that machines will continue to perform better in certain tasks than humans, especially when it comes to analysing large amounts of data.[14] However, this also makes clear that advancements in AI announced by some of the big technology companies as well as the dystopian visions about AI being a 'fundamental existential risk to human civilisation'[15] need to be carefully examined and challenged.

Understanding machine learning

One of the most important advances in AI that led to the recent, more widespread applications and impressive achievements is machine learning. Machine learning is behind AlphaGo's ability to beat human champions, Amazon's ability to make recommendations, PayPal's ability to recognise fraudulent activities, and Facebook's ability to translate posts on its site to other languages.[16]

Machine learning is one way to achieve AI, focused on giving machines the ability to learn. This is done through a combination of data and algorithms. The machines are fed data, which they 'study' using algorithms in order to discover patterns. They 'learn' from that data in order to be able to perform a task. Generally speaking, machine learning helps us to classify, cluster, and make

predictions about data. Based on this, it is able 'to detect malware, to predict hospital admissions, to check legal contracts for errors, to prevent money laundering, to identify birds from their song, to predict gene function, to discover new drugs, to predict crime and schedule police patrols appropriately, to identify the best crops to plant, to test software, and (somewhat controversially) to mark essays.'[17]

Other key terms

Algorithms are step-by-step descriptions of how to perform a task or computation, formulated in a machine language (outside of the context of AI, algorithms can also be formulated in 'natural' language). They are aimed at solving problems and, crucially, should produce the same results, given the same inputs, over and over again. There are many different types of algorithms suited to very different kinds of tasks.[18] Of particular relevance in the context of AI are machine learning algorithms and the algorithms involved in artificial neural networks. In this sense, the term algorithm, as in algorithm decision-making and algorithmic bias, and AI are sometimes used interchangeably.

Deep learning is a field of machine learning that, inspired by the structure of human brain, relies on so-called neural networks. Such machine learning neural networks can learn from data without the need for external guidance.[19] They are able to do so because they can adjust part of their programming to get better results. Deep learning is 'a terrific tool for some kinds of problems, particularly those involving perceptual classification, like recognizing syllables and objects.'[20]

Artificial neural networks power deep learning. They consist of 'a set of interconnected simulated neurons like a brain' situated in a number of layers.[21] Several layers of these neurons are simulated via algorithms. Artificial neural networks are well suited for recognising images. They are not explicitly programmed with rules to, for example, recognise images of cats. Rather, they learn from small set of images which have been labelled manually as, for example, 'cat' or 'not cat'. From this, the network identifies a set of characteristics that allow it to identify cats in images.[22] The same principle can be used to analyse satellite images, an example relevant to foreign policy.

We should also be aware that the terminology surrounding AI, such as the word intelligence itself, is ripe with words that bring an air of anthropomorphisation to the debate. The aspect of learning or training in machine learning is no different. We need to be clear that the learning which occurs as part of machine learning – for example when an AI system is increasingly able to recognise a human face in a picture or a building on a satellite image with precision – is really a matter of gaining greater accuracy in performing a specific task. Learning in this sense means that the likelihood that the system 'gets it right' has increased. This works on the basis that the system went through many more iterations of identifying specific elements in the image and made adjustments depending on whether it was successful. However, most would agree that this is very different from how we understand intelligence.

When it comes to training, we can distinguish between supervised, unsupervised, and reinforcement learning. AlphaGo was trained by playing the game against humans. Using reinforcement learning, its successor AlphaGo Zero learned by playing against itself without human input.[23,24] Reinforcement learning is now being used to train AI that is proficient at more than one board game.[25] This is an important step, given the fact that the potential application of a specialised AI is widening, at least within well-defined parameters.

Big data for machine learning: importance and challenges

We cannot talk about AI without talking about big data. AI and machine learning in particular build on an increasing availability of big data combined with increasing amounts of computational power. Machine learning is nurtured by data that is available in unprecedented volume, variety, and velocity. Big data, coupled with deep learning, is behind most of the machine learning advancements of recent years.

On the one hand, we can define big data in relation to human activity. Big data generated by what people say includes online news, social media, radio, and TV. Big data generated by what people do includes traffic movements, mobile communication, financial transactions, postal traffic, utility consumption, and emissions of various kinds. On the other hand, we can also consider different sources of big data. Digital data is generated automatically by digital services such as GPS data from mobile phones. Online information includes data generated by activities on the Internet. Geospatial data includes data from satellite and remote sensors.

The role of the Internet of Things (IoT) is one of the reasons for the unprecedented availability of big data. IoT describes the connection of physical objects, 'things', via electronics means. Devices ranging from consumer goods such as smart fridges, smart watches, and smart clothing to transportation vehicles or the sensors present throughout the infrastructure of a whole city, collect and exchange data. Terms like 'smart houses' and 'smart cities' describe the connection of physical objects via a vast array of sensors but also emphasise the importance of AI, in particular machine learning, to move from vast amounts of data to 'smart' solutions to everyday problems. In this sense, the trinity of IoT, big data, and machine learning forms one of the cornerstones of the recent advances in AI.

However, there are challenges related to the use of data for AI. Data quality is a particular concern for machine learning. In simple terms, biases in the training data will lead to biased outcomes, which is problematic if decisions with far-reaching implications are based on these outcomes. For example, the analysis of patterns and image recognition are only as good as the data that has been used to train the AI. In terms of accuracy, we need to consider whether the data is representative and captures the full extent of the phenomenon to be studied. There are many examples of biases in machine learning caused by biased data. The concern is that such biases exacerbate disadvantages along socio-economic, racial, or gender lines. An often-cited example is a version of Google image recognition, which was taught via machine learning to recognise key elements in a photo, such as recognising when a photo contained a human being. The AI tended to label photos that showed black people as containing 'gorillas', which could have been due to a lack of diversity in the training data.[26]

In addition, performing big data analysis raises questions about access to data, data interpretation, data protection, and data security. Addressed in a previous study on data and diplomacy,[27] these issues remain a key concern in the context of AI and diplomacy, particularly in cases concerning sensitive data or political decisions with potentially far-reaching consequences.

We also need to consider that dependency on big data also proves to be a limit of machine learning. It makes clear that machine learning is not applicable in areas where limited data is available. This not only limits its possibilities for application, it also highlights that machine learning is probably not the way to develop generalised AI.[28]

1.2 Mapping AI and diplomacy

To make sense of the interplay between AI as a new technology and diplomacy in broad terms, we are suggesting a three-part typology that includes AI as a topic for diplomacy, AI as a tool for diplomacy, and the AI-related shifts in the environment in which diplomacy is practised.

AI as a topic for diplomacy

The more AI becomes embedded in various aspects of our lives, the more the technology introduces shifts in established areas of foreign policy and the international agenda. As AI applications are more and more widely used, they also raise new sets of questions, many of which need to be addressed at

the international level where diplomatic efforts will remain crucial.

On the one hand, diplomats need to be able to adapt to and comfortably deal with shifts in the way existing topics are discussed due to the implications of AI. For example, the role of AI in supporting the achievement of the sustainable development goals (SDGs) has gained some prominence. Discussions include potential benefits of the use of AI in a number of areas, including increasing agricultural productivity, availing of health services in rural and remote areas, predicting and controlling the spread of diseases, enhancing sustainable marine ecosystems, and mitigating climate change.[29]

On the other hand, diplomats also need to deal with the emergence of new, AI-related topics on the international agenda. These include debates around new technology developments such as Lethal Autonomous Weapons Systems (LAWS) and debates about the ethics of and standard setting for AI.

AI as a topic on the international agenda can also be mapped by distinguishing between key themes – economy and business, security, and democracy human rights and ethics – and identifying the AI-related topics that are likely to impact these themes.[30]

- **Economy and business:** There are both disruptions and opportunities in this area. For example, there are concerns over the potential shift and concentration of economic power. At the same time, there are opportunities for leap-frogging stages of development for developing countries due to

the adoption of the technology. In this context, the potentially increasing role played by domestic and foreign technology companies is worth considering and raises questions for diplomats about how to engage with these non-traditional diplomatic actors. Lastly, labour market disruptions due to automation in various industries and jobs will have an impact on the low- and semi-skilled workforce and most likely a disproportionate impact on developing countries.

- **Security:** The balance of power between countries might shift and asymmetrical advantages could emerge with advances in the technology. The development of LAWS raises substantial concerns for international security as well as questions of ethics in warfare. Other concerns include the use of AI-powered technology in terrorism and data warfare. Increasingly, traditional methods of conflict might shift towards the cyberdomain. This also raises questions of regulation of new weapons systems, in particular LAWS.
- **Democracy, human rights, ethics:** There is potential to misuse the technology to restrict rights and liberties. In particular, this concerns discrimination due to bias in algorithmic decision-making. Existing inequalities might be exacerbated.

Chapter 4 of this report zooms in on human rights to highlight the impact of AI on one particular topic on the international agenda. Starting from the assumption that human rights apply online as they apply offline, we look at the shifts in the debate and the additional concerns raised by AI.

AI as a tool for diplomacy

Looking at AI as a tool for diplomacy aligns with the goal to '[d]evelop and promote the use of [AI] to serve humans better'.[31] From that perspective, AI is not an end in itself but rather only as useful as it can meaningfully support human endeavours – and, in the context of this report, the work of diplomats and the ministry of foreign affairs (MFA).

If we are interested in exploring AI as a tool for diplomacy, we need to start by defining what it is that diplomats do. One way of doing this is to look at the functions of diplomacy. The Vienna Convention on Diplomatic Relations describes the functions of diplomacy as information gathering, communication, representation, negotiation, and the promotion of friendly relations.[32,33,34] For our purposes, we also include the provision of

consular services as a function of diplomacy. However, to understand how AI can play a role in supporting these functions of diplomacy, we need to zoom in further, and identify very specific tasks in each.

Thinking of AI as a tool for diplomacy raises a number of interesting questions: What elements, if any, are we comfortable outsourcing to a machine? Are there any tasks that could be automated? What does 'meaningful human control', a term borrowed from the discussion on LAWS mean in this context? In what cases and to what extent do we know or need to know how an AI system arrived at a certain result? We will address these in Chapter 3 and provide concrete examples of the use of AI to perform specific tasks in diplomatic practice.

AI and a changing environment for diplomacy

AI has an impact on the environment in which diplomacy is practised. In fact, there has always been a close link between geopolitics and technology. 'Historically speaking, each epoch has its defining technology that determines economic, social, and political successes. In the past, possession of land, access to raw materials and possession of industrial technology were the defining technologies.'[35] AI might turn out to be the defining technology of our time.

AI might shift the balance of power between countries. There are concerns about a new arms race, between the USA and China in particular, for dominance in the area of AI.[36] Simultaneously, reports suggest that an era of increasing international competition might be looming.[37] New weapons systems, such as LAWS, exacerbate worries about technological advances leading to global instability and one-sided advantages for those who manage to develop the technology first. Similarly, there is the assumption that those countries that will be able to leverage the impact of AI on business and the economy first, will have a significant head start.

In this, the role of diplomats will be more important than ever. They will have to navigate between protecting national sovereignty while also engaging in meaningful co-operation.[38]

Approaching AI through the framework of science diplomacy

Another way of framing the relationship between AI and diplomacy is to build on work in the area of science diplomacy. Science diplomacy, broadly speaking, can be understood 'as a diplomatic technique by which S&T [science and technology] knowledge is freed from its rigid national and institutional enclosures, thereby releasing its potential to address directly the drivers of underdevelopment and insecurity'.[39]

In *New Frontiers in Science Diplomacy*, the British Royal Society and the American Association for the Advancement of Science outlined a three-part typology for science diplomacy: 'informing foreign policy objectives with scientific advice (science in diplomacy); facilitating international science cooperation (diplomacy for science); [and] using science cooperation to improve international relations between countries (science for diplomacy)'.[40]

This framework can also be applied to AI. Diplomacy for AI serves as a reminder that diplomacy has a role to play in supporting AI research. AI for diplomacy alludes to the possibility that international scientific co-operation in the field of AI can potentially contribute to improved relations between countries. AI in diplomacy parallels these considerations on AI as a tool for diplomacy.

Diplomacy for AI shifts the focus towards diplomatic activities that promote research and development co-operation across borders. From this perspective, diplomacy is first tasked with generating opportunities for domestic AI research centres, academic institutions, and scientists by supporting and nurturing international networks. Second, diplomacy for AI also includes the promotion of domestic research excellence abroad as part of public diplomacy. Third, diplomacy is tasked with fostering international co-operation and intergovernmental scientific institutions, such as the European Organization for Nuclear Research (CERN). In this sense, the traditional skills of diplomats, communication and negotiation in particular, are put into service to advance AI research domestically and internationally.

AI for diplomacy reminds us that scientific co-operation can act as a driver and supporter of peaceful international relations. In this regard, former US Secretary of State, Hillary Clinton, described it as 'the use of scientific collaborations among nations to address the common problems facing 21st century humanity and to build constructive international partnerships'.[41]

This approach, if it can be harnessed for AI, stands in contrast to the concerns over an AI arms race that have entered the debate recently. Collaboration that sees computer scientists, engineers, and others focus their activities on a singular goal – the advancement of AI as a research field – can promote understanding and further collaboration across borders. There are a number of examples from the Cold War, which highlight how co-operation between US and Soviet scientists was not only possible but also had an impact beyond the scientific realm. This perspective builds on the assumption that 'science is a neutral platform that allows for less politically charged dialogues, which in turn create bridges that help overall diplomatic efforts.'[42] In terms of scientific co-operation on AI across borders, the European Union (EU) offers a good example. It has taken concrete steps to ensure closer co-operation on AI, and part of its efforts focus on establishing scientific and industry links to support research and development (R&D) in the field

of AI. This co-operation is partly driven by pragmatic considerations, namely 'to increase the EU's competitiveness, attractiveness, and excellence in R&D in AI'.[43]

Yet, it is precisely this linking of research on AI across borders that could counteract arms race tendencies.

Conclusions

In this chapter, we explored some of the fundamental definitions and concepts of AI and provided suggestions for meaningfully mapping the relationship between AI and diplomacy.

Although there is no *one* definition of AI, the distinctions between generalised and specialised AI and strong and weak AI are important. AI in its current form is limited to very specialised tasks and a generalised AI does not seem to be on the horizon any time soon. Most of the recent advancements in AI are linked to machine learning coupled with the availability of big data. Algorithms, deep learning, and artificial neural networks are some of the other key terms in relation to AI.

Understanding these differences and some of the nuances as a non-technological lay person adds a layer of complexity when it comes to discussing AI and diplomacy. This calls for capacity building to ensure a basic understanding.

Talking about AI and diplomacy also highlights the need to mediate between two very different worlds when discussing the interplay between the two: the social and political world of the diplomat and the technical world of the computer scientist and technologist. An important starting point is for diplomats and policymakers to develop a nuanced understanding that is grounded in examples of current applications of AI.

With regard to the relationship between AI and diplomacy, it is not enough to point out that diplomats will encounter new issues due to the increasing role played by AI in various fields. Doing so means to miss important aspects of the way AI and diplomacy relate to each other. To offer a comprehensive as well as a concrete understanding of the relationship between AI and diplomacy, we suggest a distinction between AI as a diplomatic topic, AI as a diplomatic tool, and AI as a factor that reshapes the environment in which diplomacy is practised (in particular in geopolitical and geo-economic terms).

In addition, we suggest applying a typology developed in the area of science diplomacy and thus using diplomacy for AI, AI for diplomacy, and AI in diplomacy as an additional framework. The important contribution of this typology is located in the fact that it serves as a reminder that (a) diplomacy plays a role in supporting the development and application of AI (diplomacy for AI) and (b) that AI, understood as the co-operation between researchers and developers across borders who work towards a common goal, can contribute to peaceful international relations (AI for diplomacy). While this inception study is not the place to flesh out these typologies in detail, we will return to them as ways of framing the debate in subsequent sections of this report.

2. National AI strategies and international co-operation

Governments and international organisations have begun to think extensively about artificial intelligence (AI) and its economic and social implications. This is reflected in the preparation and publication of national AI strategies and reports and the formation of new government bodies and commissions.[44] Looking at these developments not only gives a sense of the importance accorded to the topic but also the variety of responses and priorities.

Priorities set in national AI strategies and reports are also relevant for the work of diplomats, who have the role of supporting the achievement of national strategies – for example by encouraging co-operation, facilitating dialogue between national and international companies, and supporting global discussions on norms and ethics related to AI. For those working in the field of international relations, national AI strategies offer a first point of orientation with regard to the kinds of priorities that countries will also articulate and pursue internationally.

In addition to national strategies, international organisations, in particular the European Union (EU) and United Nations (UN) system, have taken up AI as a topic. While these developments are still very much in their infancy, we can already get a sense of the priorities discussed in these venues.

Given these developments, this chapter provides an overview of and commentary on the various strategies and reports, in particular zooming in questions surrounding (big) data for AI. It addresses questions of competition and co-operation between countries and ends with reflections on the next steps for ministries of foreign affairs (MFAs) in this context.

2.1 Overview of national AI strategies and related documents

In recent years, many governments have published national AI strategies or reports designed to lead to or prepare such strategies. Some governments were keen to be among the first to develop them in order to be seen as early adopters and to highlight their awareness and engagement with the topic. This is associated with the idea that there is a 'significant competitive advantages for early movers'[45] in the field of AI. The rapid succession of publications of national AI strategies over the last two years, since 2017, was described as 'the race to become the global leader in artificial intelligence'.[46] While this competitive perspective is definitely not reflected in all of the strategies, many of them aim at leadership in the field.

Indeed, the emphasis on the need for countries to remain competitive or advance towards playing a *leading* role in the area of AI is a very prominent feature of many of the documents analysed. Through the strategies and reports, governments have, for example, expressed the desire to

- be 'at the forefront of the artificial intelligence and data revolution'.[47]
- harness 'the opportunities that the use of AI can offer, with the aim of strengthening […] welfare and competitiveness'.[48]
- keep 'a competitive edge in developing machine learning and vision, natural language processing and other key AI fields'.[49]
- work towards 'achieving and maintaining leading global excellence in the research, development and application of AI'.[50]

China's strategy sets three very ambitious, but also very concrete, goals in this regard: First, reaching a level of AI development that is comparable with globally advanced levels by 2020; second, achieving major breakthroughs in basic theories of AI by 2025; and third, being a world leader in AI theories, technologies, and applications by 2030.[51]

In terms of content, even a cursory look at the strategies and reports shows that the majority of them approaches AI with a very positive outlook, often associated with the potential for economic growth and greater efficiency in many sectors across the economy. The national strategies and reports also mention challenges, but overall, the opportunities and benefits associated with AI outweigh these.

Analysing the strategies and related documents offers a starting point to highlight differences and similarities between countries and to discuss the implications of these strategies for diplomacy, both in terms of international co-operation and international competition. However, it is worth emphasising that the strategies and reports are diverse and only comparable up to a point. They originate from different ministries, government entities, or other bodies, which leads to a different focus and potentially different target audiences. In addition, it is worth noting that some countries incorporate their goals for the development and application of AI under broader digital strategies, such as Belgium, Denmark, Italy, Norway, and Switzerland.[52]

Some strategies and reports are very comprehensive, while others focus only on specific aspects of AI. For example, Canada, one of the first countries to release a national strategy, zooms in on supporting research, training, and innovation in AI in its strategy.[53] Italy's white paper focuses on public administration service provisions. Having said this, however, the absence of a specific topic from a national strategy or related report does not mean that it is also absent from the entirety of what a government plans and does in the area of AI.

The role of data

Data, and in particular big data, is crucial for the digital economy and for developing AI applications. Hence, access to data and the way it is governed become important questions for national AI strategies. In the following, we give an overview of how different countries address questions surrounding data in their national strategies and related documents. All of the documents mention data but the degree to which countries engage with its governance in their national AI strategies varies.

In its 2017 innovation roadmap, Australia encourages an increased use of open data by companies and stresses the importance of making data publicly available for use by businesses in order to increase productivity.

When it comes to national AI strategies, the USA is an interesting case. While three papers from 2016, prepared by the Office of the President and the National Science and Technology Council, set a national discussion in motion, no national strategy has been developed and no additional documents have been published since. Internationally, the USA, and Silicon Valley companies in particular, is at the forefront of innovation in AI. Private sector actors, especially those located in the USA, generally emphasise that in order to stay globally competitive and innovative, regulations on the sector should be limited. Yet, it is also recognised that when it comes to building and supporting a national ecosystem, governments play an important role in (a) supporting the development and adoption of AI applications across various sectors of the economy, (b) ensuring the development of relevant skills in the workforce, (c) providing reasonable access to (big) data, and (d) providing publicly funded research and development (R&D) opportunities.[54] Hence, some observers argue that the absence of a national AI strategy might endanger the USA's leading role in AI in the medium and long term.[55] More generally speaking, this discussion on the importance of developing national strategies serves as a reminder that those countries that do not have such a strategy yet, need to focus on developing one that is appropriate to their specific context.

A close and comparative analysis of each of the national strategies that has been published so far is beyond the scope of this chapter. Yet, we look at one topic of great relevance for AI in more detail. We ask: How do the strategies and related reports discuss the role of data and what measures are they suggesting to ensure that (big) data on which the development of many AI applications depend is available and easy to access?

The roadmap also identifies risks related to privacy, de-identification, and security and argues that it is important to 'balance the need to instil trust and acceptance of data systems within the community with the need to empower citizens, governments, industries, and researchers'.[56]

China places prominence on big data for research breakthrough and describes this as a key factor in economic growth. It aims to build and rely on 'a national data sharing exchange platform, [an] open data platform and other public infrastructure'.[57] The strategy also points out that the vast amount of data available to the Chinese government and businesses lead to a competitive advantage.

Finland's interim report argues that the country's strength lies in the availability of high-quality data, which makes the country attractive for companies and researchers. In comparison to AI development, application, and investment in the USA and China, 'Europe is very clearly lagging behind [though catching up], and development is hindered by such things as heterogeneous legislation and data resources.'[58] In addition, the report stresses the need to strike a balance between protecting sensitive data and individuals' rights and the business opportunities for public service provision offered by that data. Ensuring greater data interoperability is mentioned as another factor in making AI work on a broad basis.[59]

Highlighting the centrality of data when it comes to AI, the first chapter of the Villani report, the central document to understand France's national AI strategy, is dedicated to data.

France starts from a position that is critical of the status quo when it comes to data. The report states:

> ' Data currently mostly benefit just a handful of very large operators, so greater data access and circulation will be required to restore a more even balance of power by extending these benefits to government authorities, as well as smaller economic actors and public research.'[60]

In response to the data needs of AI, the Villani report argues that new means of sharing, governing, and producing data need to be developed. It emphasises that data is a common good and that economic actors need to be encouraged to share their data. The report makes another strong statement. It declares data an issue of sovereignty and argues: 'It is vital for France and Europe to maintain a firm stance on data transfer outside the European Union.'[61]

One of the main goals of the German strategy points out that the country has a 'specific data stock', which should be used to develop AI-based business models in Germany to make the country become a 'new top export[er], whilst strictly observing data security and people's right to control their personal data'.[62] Data sharing should be made easier to promote co-operation between business and research institutions and a national research data infrastructure should be built to enable centralised access for researchers.

In its discussion paper for a national strategy on AI, India describes the 'absence of an enabling data ecosystem' as a barrier to using AI at scale. The paper also argues that privacy and security are important, but that there is currently a lack of formal regulation around the anonymisation of data.[63]

Italy's white paper identifies various challenges around data as one of the reasons for the lack of adoption of AI applications. This concerns data quality, potential biases of data, the role of open data held by public bodies, and the possibility of equal and non-discriminatory access.[64]

Japan sees room for improvement in the lack of data digitalisation some areas. However, it also stresses that not all AI applications are reliant on big data and encourages the exploration of 'high-precision learning from small amounts of data'.[65]

Mexico's white paper suggests investing in infrastructure to provide good quality data, in particular to build a 'resilient open data infrastructure', while taking care of protecting privacy.[66]

New Zealand does not have a fully developed AI strategy yet but has published its *Analysis of the Potential Impact and Opportunity of Artificial Intelligence on New Zealand's Society and Economy*. The report suggests increasing data availability and accessibility, releasing public data, publishing localised data repositories, and establishing data trusts, which are 'frameworks and agreements to ensure the safe, trusted and efficient exchange of data between public and private sector organisations'[67] to enable AI-based solutions. In this context, localised data repositories will 'improve the accessibility of local training data sets that are uniquely Kiwi, instead of relying on overseas data sets'.[68]

The 'Declaration on AI in the Nordic-Baltic region', a joint declaration by Denmark, Estonia, Finland, the Faroe Islands, Iceland, Latvia, Lithuania, Norway, Sweden, and the Åland Islands, states that data, along with infrastructure, hardware, and software, should be based on standards that enable 'interoperability, privacy, security, trust, good usability, and portability'.[69]

South Korea has the goal to establish 'a national data management system for the development of largescale data infrastructure that facilitates machine learning'. It also suggests that data owned by the government should be in open formats.[70]

Sweden stresses the importance of access to data and infrastructure, including the computational capacity, and emphasises that it has an 'almost unique volume of high quality data, which is often a prerequisite for AI applications and if properly managed can contribute to creating considerable benefits'.[71]

The *AI in the UK: Ready, Willing and Able* report points to challenges experienced by small and medium enterprises in accessing data, especially when compared to US-owned technology companies who have greater opportunity to either purchase data or are large enough to have generated the needed data on their own. It also states that the UK is still in need of a strategy to actively create big data, especially in areas of government interest such as healthcare, transport, science, and education.[72] The report suggests stimulating 'access to data'[73] by making public datasets open and available, upgrading data infrastructure, which is 'critical to leading the world in AI'[74] and identifying barriers to data sharing.

A few countries also mention the General Data Protection Regulation (GDPR), the EU's data protection and privacy regulation, when discussing data in the context of AI, and automated decision-making processes that use personal data in particular. The strategies of Germany and France specifically emphasise the need to develop and use AI systems 'in a way that is compatible with data protection rules'.[75] As part of the overall goal of adapting regulatory frameworks in light of AI systems, the German strategy suggests assessing 'how AI systems can be made transparent, predictable and verifiable so as to effectively prevent distortion, discrimination, manipulation and other forms of improper use, particularly when it comes to using algorithm-based prognosis and decision-making applications'.[76] The provisions of the GDPR form a cornerstone of this. Similarly, Sweden states that the data protection principles offered by the GDPR make the regulation an important part of any AI framework.[77] New Zealand cautions that the ability to comply with the GDPR will depend on progress with Explainable AI, AI that documents and explains how it arrived at decisions, and the uptake of such techniques.

To succeed in adapting to the era of AI, countries, generally speaking, place a strong emphasis on open data and the important role played by appropriate data repositories and data infrastructure. Governments commit to providing more open data and encourage the private sector to do the same. Some stress the importance of small data, especially in the absence of easily accessible big data in a relevant field, and the development of AI tools that work with small data. Some countries mention the importance of unique, localised, or small data and argue that the uniqueness or quality of the data they hold will give them a competitive advantage.

2.2 International co-operation

A number of the national strategies, especially the European ones, emphasise the need for co-operation across borders. While co-operation is particularly important for smaller countries, very few countries can actually do it alone when it comes to fully harnessing innovation and economic benefits based on AI technology that is driven by big data. For most countries, the national AI strategy has to be put in an international context as their national AI ecosystems are too small. Observers have argued that even countries like Germany, being the fourth largest economy in the world, depend on co-operation in order to be competitive internationally.[78] The USA and China are the only exceptions to this rule, given the amount of data the two governments and their digital industries have at their disposal.

National strategies and international co-operation

Some of the national strategies analysed here explicitly stress international co-operation. The Chinese strategy states the country will 'actively participate in global governance of AI, strengthen the study of major international common problems such as robot alienation and safety supervision, deepen international co-operation on AI laws and regulations, international rules and so on, and jointly cope with global challenges'.[79] The Finnish interim report notes that 'strong international co-operation is a prerequisite for Finland's success' in an environment of global and strong competition.[80] Co-operation is not only envisioned between governments, but between other stakeholders as well, including through public-private partnerships. France and Germany plan to co-operate through a Franco-German research and development network. The French national strategy describes this co-operation as a factor in the development of a European industrial policy on AI. It adds that Italy, due to its advances in robotics, should be seen as a 'possible serious partner' in this endeavour as well.[81] The

report drafted in preparation for the Mexican AI strategy argues that companies will benefit from a coherent operational framework, which requires countries to co-operate in order to develop such a framework for AI.[82] The declaration of the Nordic-Baltic region includes a commitment of the countries of the region towards greater co-operation. In addition, it also includes a commitment to strengthen the role of the EU and support European discussions and initiatives.[83] Sweden mentions co-operation on AI research with other countries.[84]

In short, looking at the strategies analysed here, drivers for international co-operation in the area of AI include
- the need to pool resources for research and development.
- the need for access to (big) data.
- a desire to set standards and develop an ethical framework.
- a desire to build on and amplify shared norms and ethics.

Co-operation at EU level

In April 2018, 25 European countries signed a declaration of co-operation on AI, in which they committed to working together towards 'increasing the EU's competitiveness, attractiveness and excellence in research and development in AI'.[85] They also agreed to address the social, economic, legal, and ethical concerns related to AI and to prevent the development of harmful AI.

The European Commission's strategic note, *The Age of Artificial Intelligence*, makes some concrete suggestions regarding co-operation at the EU level. The note argues that, in comparison to the USA and China, the EU is lagging behind when it comes to the adoption of AI. In order to shift this, the EU should build on two approaches – '[F]irst, creating an enabling framework favouring investment in AI, and second, setting global AI quality standards'.[86]

Looking towards supporting AI development and uptake in Europe, the strategic note suggests[87]
- making access to data easier.
- enabling infrastructure investment in areas of importance for AI, such as high-performance computing.
- promoting the development of AI hubs and support research excellence,
- creating a permanent network of AI research institutions.
- supporting the creation of a European Artificial Intelligence Platform to bring together different stakeholders from various sectors in order to identify areas in which the AI ecosystem in Europe need strengthening.

The strategic note also argues that the EU should become more active in multilateral forums 'to take the lead in an international multilateral discussion around the use of Artificial Intelligence for military purposes, and to promote global solutions, including blanket bans'.[88]

Like some of the European countries mentioned, the strategic note describes the GDPR, which had yet to come into force when the note was published, as a competitive advantage for AI developments in Europe. The GDPR ensures data quality, which is essential for applications of AI.[89]

AI at the UN

The UN has also begun addressing questions of AI. The role of AI in supporting the implementation of the sustainable development goals (SDGs) is just one example of why AI has also become relevant for the UN.

Two UN strategies are worth noting: *The UN's Secretary-General's Strategy on New Technologies* and the *System-wide Strategic Approach and Roadmap for Supporting Capacity Development on Artificial Intelligence*.

Although not strictly focused on AI, the strategy on new technologies sets important goals for the organisation as a whole. These include 'deepening the UN's internal capacities and exposure to new technologies; increasing understanding, advocacy and dialogue; supporting dialogue on normative and cooperation frameworks; [and] enhancing UN system support to government capacity development'.[90]

The UN's *System-wide Strategic Approach and Roadmap for Supporting Capacity Development on Artificial Intelligence* is not only more specifically focused on AI, it also zooms in on questions of capacity development. The report acknowledges the role of AI in achieving the SDGs. It aims to ultimately work towards a guide for UN agencies in AI-related action and support UN member states in their use of AI and in reaping the benefits of the technology in developing and least developed countries in particular. The report also mentions that least developed countries (LDCs), landlocked developing countries (LLDCs), and small island developing states (SIDS) are the countries with the most need for support in developing AI-related capacities. It argues that AI capacity building programmes should not only be multilateral but also multistakeholder. All in all, the report stresses that capacity development efforts are needed in order to leave no one behind in the era of AI.

The report identifies four key areas where support is needed:
- Infrastructure: There is an increased need for Internet infrastructure.
- Data: There is a need for more robust, open, inclusive, and representative data sets.
- Individual capacities: There is a need to create more AI experts in both developed and developing countries. Particular attention should be paid to recruiting women and girls to study subjects such as science, technology, engineering, and mathematics.
- Policy, legal, and human rights frameworks: Many of the existing frameworks are not suited to respond to developments in AI. In particular, states need assistance to ensure that advances in AI do not harm vulnerable populations

The report was further discussed at the High-level Committee on Programmes (HLCP) of the UN System Chief Executives Board for Coordination (CEB). In an effort to further narrow the focus of the UN systems activities, the HLCP suggested to focus on 'AI capacity-building for developing countries, around specific priority needs such as the digital divide, infrastructure, AI governance, and human rights, with a distinct focus on reaching the bottom billion'.[91] Hence, capacity building is an area of international co-operation for AI that is likely to warrant more attention in the coming years.

The suggestion of an IPCC for AI
Speaking at the 2018 Internet Governance Forum (IGF), French president Emmanuel Macron suggested the creation of an international body for AI that is modelled on the Intergovernmental Panel on Climate Change (IPCC). He highlighted Canada as a key partner and argued that France's G7 presidency would offer the opportunity to push the idea forward. The IPCC for AI would focus on the ethical, technical, and scientific dimensions of AI and work with civil society, top scientists, and all the innovators and be supported by international organisations such as the Organisation for Economic Co-operation and Development (OECD) and the United Nations Educational, Scientific and Cultural Organisation (UNESCO).[92]
The IPCC focuses on providing scientific assessments on climate change and its implications to policymakers. Scientists work, on a voluntary basis, on regular assessment reports that review the existing literature and current state of knowledge on climate change.
Macron's suggestion has raised concerns about a doubling of efforts and attempts at global governance for AI that could harm innovation.[93] Yet, assuming that the IPCC for AI follows the principles of the original IPCC, it would bring

scientists and practitioners from all around the world together to create assessment reports that offer a full picture of the state of AI when it comes to its use and research. Based on the assessment report, government representatives negotiate a final document, the Executive Summary, that reflects the key points and, in the views of negotiators provides a balanced picture. The resulting document reflects the global consensus, based on the most recent scientific findings, on the issue.

Conclusions

The strategies and reports analysed here show that AI is on the agenda for many governments. It is treated as a cross-sectoral issue with many implications in a wide range of areas. Progress towards the goals set in national strategies and related documents depends on a number of factors. Advancing AI will happen in a complex system of overlapping spheres of economy, society, politics, and technology and innovation. It will require the co-ordination and co-operation of a diverse set of actors. Thus, AI presents a complex challenge for diplomacy.

Comparing national AI strategies and related reports with regard to their view on (big) data, we find that access to data (including questions of open data and the differences between countries regarding their ability to draw on large data sets), and building or maintaining appropriate data repositories and data infrastructures are some of the key concerns of countries as they look towards economic growth and prosperity on the basis of increased application of AI. In contrast to the assumption that AI always requires big data, some countries emphasise the quality or uniqueness of the data they hold as a competitive advantage.

A number of countries emphasise the importance of international co-operation. Their main reasons for doing so include the need to pool resources for research and development, the need for access to (big) data, a desire to set standards and develop an ethical framework, and a desire to build on and amplify shared norms and ethics. The EU focuses on closer co-operation and co-ordination in a variety of forms. It also aims to set global quality standards for AI applications and to become more active in multilateral discussions, especially on the use of AI for military purposes. The UN approaches discussions around AI with suggestions for internal capacity development. In addition, the organisation is concerned about a widening gap between developed and developing countries, especially given the potential of AI to support the implementation of the SDGs.

3. AI as a tool for diplomacy

Thinking about artificial intelligence can help clarify what makes us human – for better and for worse.[94]

As we are now about to enter a more sophisticated phase of the Digital Age, it is important to ask ourselves how the emerging digital technologies may shape the next stage of digital diplomacy.[95]

It takes time to adapt to new technologies and integrate them in existing workflows. Once a technology has become widespread and pervasive however, we often do not even notice it anymore in our daily work. Some of the artificial intelligence (AI) tools discussed here, such as tools that can analyse vast amounts of textual data, could soon be on their way to becoming exactly that: pervasive to the extent that they become invisible. In this context, it is useful to remind ourselves that e-mail was once considered a novel form of communication. Will AI tools eventually follow the telegraph, the telephone, and e-mail and become seamlessly integrated into the everyday working lives of diplomats? For example, in the mid-1990s, there was still speculation as to the extent e-mail would be taken up by ministries of foreign affairs (MFAs). It was argued that 'a new IT-influenced diplomatic interaction symbolised by the exchange of communication through dedicated e-mail'[96] could emerge as more and more MFAs begin to use information technology more widely.

Looking at AI and diplomacy means that questions regarding human-machine interactions become more prominent. We already live in the age of close interaction between humans and machines driven by AI. One only needs to think of the use of robots in production lines and warehouses where humans and machines work side by side. Of course, the importance of this interaction is not limited to machines that exist physically (robots); it also includes virtual assistants, voice recognition, and planning and optimisation algorithms. While there are discussions on the security and military applications of AI decision-making systems,[97,98] AI as a tool for diplomacy has received very little attention so far.[99]

To address questions of AI as a tool for diplomacy, this chapter first introduces the distinction between assisted, augmented, and automated intelligence and discusses one example from diplomatic practice at the borderline of assisted and augmented intelligence. Second, since developing AI applications for diplomacy will, in most cases, involve working with the private sector over time, we add reflections and recommendations regarding this process.

3.1 Possibilities between assisted, augmented, and automated intelligence

As a starting point, we build on the distinction between assisted, augmented, and automated intelligence.[100] Assisted intelligence supports the work of a human being; augmented intelligence allows humans to do something that they would otherwise not be able to accomplish; and automated intelligence describes those cases in which the entire task is performed by an AI. While automation receives lots of attention, partially due to concerns about the loss of jobs, it is safe to say that, at this point in the development of AI, 'the technology's greater power is in complementing and *augmenting* human capabilities.'[101] A 2018 report by PricewaterhouseCoopers argues that currently, the best applications for AI in business are (a) the automation of simple tasks and processes and (b) the analysis of unstructured data.[102] This is also applicable to the world of diplomacy and offers a very pragmatic approach to exploring AI as a tool for diplomacy that avoids raising unrealistic expectations.

Assisted and augmented intelligence for diplomacy

The practice of diplomacy is very closely intertwined with the written word. Diplomatic reports and treaties form a cornerstone of the profession. Analysing these reports and treaties, for example in preparation for negotiations, can be a time-consuming task for diplomats. Traditional computerised search methods are not helpful in this instance as the textual data is mostly in unstructured form, meaning that it is not organised in a predefined manner and not annotated in a way that would make it readable by a computer.

In this regard, AI can be a useful tool in facilitating tasks that require diplomats to work with, digest, or research vast amounts of text. The natural language processing (NLP) methods[103] involved in this not only allow for dealing with unstructured data, they also have the potential to uncover new connections between issues. The particular example we look at here comes from the realm of trade negotiations, which is characterised by a plethora of particularly lengthy agreements.

The Cognitive Trade Advisor (CTA), a piece of software designed to support diplomats in preparing for international trade negotiations, was launched at the 2018 World Trade Organization (WTO) Public Forum.[104] Developed based on a co-operation between the International Chamber of Commerce and IBM Watson, the system answers questions on rules of origin across a vast number of trade agreements. Rules of origin define where a product originated and consequently what duties and/or restrictions apply in cross-border trade. Trade negotiations in this area tend to be very complex and require extensive preparation, partly due to the number and length of relevant agreements.

In the process of developing the tool, the first step was to define its scope, the focus being on rules of origin in this case. The next and very crucial step was to train the CTA. Two diplomats classified rules of origin in a form understandable for the software and annotated trade agreements and relevant products to provide a basis from which the CTA could learn to recognise rules of origin across all documents and correlate them with relevant products.[105] This was an important step in order to build a basis from which the software could begin to recognise rules of origin and respond to complex queries. The CTA also includes a virtual assistant that can answer questions on the rules of origin and provide graphical illustrations of search results, thus allowing for a more natural interaction between the

tool and the diplomats and for easy comprehension of a complex issue.

Those involved in the development of the CTA, described its advantage as 'tackling the challenge of complexity'[106] when it comes to making sense of a complex issue across large numbers of treaties. They also expressed the hope that, ultimately, the insights generated by the CTA could lead to the development of a 'new model for rules of origin in trade treaties'.[107] Further, they argued that the tool could be important to support negotiators from smaller negotiations teams and level the playing field between small and large delegations. Their suggestion was to ultimately make such tools freely available, through working with United National (UN) bodies such as United Nations Conference on Trade and Development (UNCTAD).

It is important to keep in mind that the CTA is a highly specialised tool, trained for a very specific purpose. This means that for different kinds of negotiations, even different aspects of trade negotiations, a new tool has to be trained. Having said this, simpler tools that can be used to analyse social media or aggregate news from various sources are already available and used by some MFAs. This can take the form of using off-the-shelf software applications offered by technological companies or working on customisation of the tool for the specific purposes of the ministry.

In addition, MFAs have become increasingly interested in AI tools to monitor open data for early crisis detection – with the aim of expanding this towards gaining insights for crisis *prevention*. The Open Source Unit of the UK Foreign and Commonwealth Office was an early adopter in this regard. Recently, the Department of Crisis Prevention, Stabilisation, Post-conflict Care and Humanitarian Aid of the German Federal Foreign Office launched a comparable initiative that is 'evaluat[ing] publicly available data on social, economic, and political developments' in order to detect crisis at an early stage.[108]

Other examples of AI tools used to support the work of diplomats include the use of machine learning to evaluate satellite images to recognise patterns of, for example, refugee movements. More complex systems that combine insight from satellite images, textual data, and other data sources to make informed recommendations are also possible.

3.2 Approaching the development of an AI solution

In contrast to the technology, security, and health care industries, diplomatic institutions are more likely to be in a position in which they are yet to embark on developing AI solutions.

In this section, we share points for consideration for those ministries considering such a step. This builds on experiences and lessons learned from DiploFoundation's Data Team and AI Lab in their search for a software solution that responds best to established project requirements. In doing so, we map out five main steps that should be considered when developing a project that requires an AI software solution. In addition to considering these steps, it is also crucial that the MFA team working on the AI solution remains flexible and openminded throughout the process, primarily because, over time, the project might take on a different course of action.

Step 1: Determine needs, and set goals and key performance indicators

As with everything related to complex analytics, the key step is to determine whether an advanced software solution that is capable of analysing large quantities of data is actually needed. If so, it is important to clearly establish what it is that is to be achieved or discovered.

It is important to be very specific about the goals. To begin with, it is crucial to identify needs.

Questions such as the following are very useful to ask this stage:
- Is there a need to analyse historical records and identify patterns – an area of descriptive analytics?
- Is this about performing sentiment analysis on documents, social media communication, or even speeches?
- Is the important aspect the ability to analyse large amounts of documents according to specific criteria?
- Is the interest located in forecasting future events as part of predictive analytics?

In most cases, this is where many organisations fail – lack of proper (business) requirements and clear-cut ideas of what they want to achieve can result in overspending, impractical and inadequate solutions, as well as significant loss of time and resources.

It is therefore important to take the time to think thoroughly about what the MFA wants to achieve. For instance, for large quantities of data such as social media data used for sentiment analysis or key topic identification, an automated solution significantly shortens the time needed for this task. The same applies to identifying patterns and seldom-spotted information in government policies and legal documents, as we discussed using the example of the CTA.

After determining what the MFA wants to achieve, the next step is to set goals, i.e., define what is to be measured and how – units, performance, data representation, and visualisation. Raw data *per se* is useless without a context-specific situation and adequate presentation. Data and visualisations bring order to chaos and help an average user understand even the most complex data.

Step 2: Create a project description

The next step is to draft a project description with clearly specified requirements, identified tasks, and clear objectives. At this stage, the purpose of the project or research needs to be clearly identified. Whilst creating the project descriptor, it is imperative to consult with different stakeholders and departments within the MFA in order to build a wider picture of the requirements.

Furthermore, a technical adviser should be involved in the drafting process so as to assist with technical demands that diplomats and other staff might not be familiar with. This exercise is crucial as it allows the team engaged in the task to get a clear picture of what has to be done and in what way.

Step 3: Research

Once the project description has been created, the research phase can begin by analysing the market and technical capabilities online. Existing solutions can be explored and demos can be tested based on the needs and goals. Even though the ultimate objective is to obtain a software solution that relies on technological strengths (and reduces workload), it is still recommended to test and compare the obtained results manually to see whether the software gives correct outcomes. The research phase is very useful because in certain situations expectations might not correspond to reality and consequently the obtained results might not be suited to the given objectives. In other words, this

phase is about exploring what current AI technologies can and cannot deliver.

At this stage, advantages and shortcomings of the different software solutions that are tested can clearly be identified. If shortcomings outweigh advantages, they might be addressed through slight modifications to the existing program. This approach of slight modification of existing solutions is particularly useful if the goal is to develop a small pilot project to showcase the potential of AI applications. However, if modification is not desirable or not possible, a customised solution that corresponds best to the need of the particular project should be considered. Customising solutions builds on close co-operation between MFAs and software companies. Hence it is important to think about the availability of personnel to engage with technological companies.

Step 4: Negotiate

When it comes to defining the relationship with software developers and technological companies, there are two main strategies.

A strictly professional-client relationship can be adopted, when it is enough to obtain the required software solution. This entails purchasing the service and receiving the necessary technical assistance and support in return. This option probably does not involve

substantial negotiations. It also gives a lot of independence and perhaps requires the least effort. However, following this approach means forgoing opportunities for co-operation and customisation.

An alternative to this strategy is to seek a partnership in the form of a mutually beneficial relationship for the business and the ministry. In this case, it is crucial to carefully define the parameters of the partnership. Parameters are, among other factors, determined by the defined scope of the project, the financial resources available, and the kind of data utilised. If the project deals with sensitive data, in contrast to openly available data, the terms of the partnership need even closer attention, in particular with regard to data ownership and confidentiality.

Step 5: Implement

At this stage, it is important to ensure that implementation functions smoothly and that the means necessary for successful completion are present. Additional internal measures might be necessary such as acquisition of licences, server updates, and staff support. At the same time, monitoring and evaluating are important in order to determine whether
- expectations are met.
- there is a return on investment.
- the co-operation is beneficial for the MFA.

Conclusions

AI not only impacts the kinds of topics diplomats need to address, it also adds to the arsenal of tools at the diplomat's disposal. In order to think through AI's potential to serve as a tool for diplomats, the distinction between assisted, augmented, and automated intelligence is a useful starting point.

At this point in the development of AI, it is best used to assist and augment intelligence. Since a substantial part of diplomatic practice is concerned with textual data, for example in the form of treaties and diplomatic reports, tools that can support the meaningful analysis of this data at scale are of particular interest.

In this sense, AI that uses NLP techniques to make sense of large amounts of unstructured data in natural language provides some of the greatest promise when it

comes to AI as a tool for diplomacy. The analysis of texts at scale has the potential to make the work of diplomats more effective and free up time and resources. As a consequence, more time can be spent on aspects of diplomatic work that require uniquely human skills and human intuition.

Small and developing countries might struggle to develop these tools on their own. But if shared or publicly available, AI tools, similar to the one discussed in this chapter, can make a substantial contribution to levelling the playing field at international negotiation tables.

Lastly, MFAs need to be prepared to engage with the business sector to obtain these tools and in most cases of more complex applications, close co-operation over time is needed.

4. Addressing the human rights dimension of AI

The main risk today is the gap between innovations and our legal framework, whose basic concepts are not suited for the virtual world.[109]

Machines function on the basis of what humans tell them. If a system is fed with human biases (conscious or unconscious), the result will inevitably be biased.[110]

In this chapter, we address the human rights dimension of AI to highlight the impact of new technology on established topics on the international agenda. We look at the topic from a state-centred perspective to highlight the importance of diplomats and other officials in addressing these issues while being keenly aware of new concerns arising due to the impact of AI. This also highlights potential areas in which additional capacity building will be needed in order for diplomats and other officials to be effective in addressing the human rights dimension of AI.

The protection of human rights and fundamental freedoms is a central part of the foundational contract of modern democratic states, forming part of their raison d'être. Their legitimacy to act depends on putting people at the centre of their ambitions as evidenced by international legal texts, including the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and other international and regional instruments. These instruments and frameworks were established following the atrocities of two world wars which asserted the need to protect everyone from abuse and harm, in particular underscoring the dignity and integrity of every individual. Ever since, individual civil, political, economic, social, and cultural rights and freedoms have been interpreted and developed by courts and international organisations around the world.

States are the duty bearers for the protection of human rights. They are its primary guarantors and, under certain circumstances, they can be held accountable. They have both positive and negative obligations to protect and to refrain from interfering with the exercise and enjoyment of rights and freedoms. In many ways, the Internet and information communication technologies have enabled the exercise and enjoyment of many rights and freedoms while, at the same time, they have challenged certain among them, due to emerging problems, such as online hate speech, incitement to violence, and fake news. The role of the state is to strike a balance between freedom and protection, rights and responsibilities. This balance also requires that the state makes efforts to engage the private sector in respecting human rights, including in the design, development, and delivery of their digital services.

4.1 Human rights in the digital era

To begin this discussion, it is crucial to recognise that human rights apply online as they do offline. For the past 15 years, we have seen a growing concern about this question, in particular at the international level.

In 2003–2005, the United Nations' World Summit on the Information Society (WSIS) resulted in four outcome documents which paved the way for multistakeholder dialogue on *inter alia* finding 'solutions to the issues arising from the use and misuse of the Internet, of particular concern to everyday users'.[111] In 2012, the UN Human Rights Council (HRC) 'affirm(ed) that the same rights that people have offline must also be protected online'.[112]

In 2014, the HRC underlined the need to address the issue of security and human rights together, coupled with support for access to a global and open Internet.[113] In 2016, the HRC stressed *inter alia* the importance of not disrupting access to the Internet and the need to combat online hate speech which constitutes incitement to discrimination or violence.[114] This has been complemented by reports, opinions, statements, and instruments of UN special rapporteurs, the High Commissioner for Human Rights, and the work of regional organisations such as the Organisation for Economic Co-operation and Development (OECD), the European Union (EU), and the Council of Europe.

The courts continue to articulate the need for narrowly circumscribed limitations to rights and freedoms to protect national security, to fight crime and disorder, and to respect the rights of others. To this end, they have interpreted existing laws and filled gaps when norms do not exist, constituting de facto makers of digital rules.[115] There have been landmark cases on privacy and data protection, freedom of expression, and intermediary liability to name but a few, which have given rise to certain precedents. Complemented by intergovernmental frameworks, they continue to address the challenges of

technology in their judgments and opinions. Their iterations have provided checks and balances to governmental activity, strengthening the resolve of states to protect and foster respect for human rights online while providing guidance when national laws do not comply with international human rights frameworks.[116] In this sense, courts have emerged as active makers of digital policy.[117]

Human rights, which are universal, indivisible, and interdependent, prevail over contractual clauses, such as general terms and conditions, imposed on Internet users by the private sector.[118] They are broad in scope and cover an array of rights and freedoms. They exist to ensure that the highest attention is paid to the lives and dignity of everyone, to ensure that no one is left behind. States are the primary guarantors of these rights and freedoms and are formally obligated to protect them, whereas companies, which drive the development and deployment of AI-enabled products and services, have no such obligation. Not bound by international law, their commitment is voluntary. Consequently, it is the duty of states to encourage such companies to design, develop, and deploy AI with respect for human rights above and beyond the ethical commitments the private sector might undertake.

4.2 Specific rights in relation to human rights

Having said this, it is crucial to look at specific rights and how they are (potentially) impacted by AI. In the following we look at the right to privacy, family, home, and correspondence; the prohibition of discrimination; and the right to hold opinions and to freedom of expression.

The right to respect for private and family life, home, and correspondence

The ICCPR states that '[n]o one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honor and reputation.'[119] The right enables the individual control of personal information online and offline, and includes the decision whether to disclose it. In other words, the individual has the freedom to be left alone in order to develop their own personality, to live with dignity and in security. The right is not absolute and can be subject to narrowly circumscribed limitations in order to protect other legitimate rights and interests such as national security, fighting crime and disorder, etc. Originally prescribed to safeguard against arbitrary interferences by public authorities, this right also encompasses interferences by companies and by individuals. These interferences include the publication of personal information; the security and privacy of mail, telephone, e-mail, and other forms of

communication, including online information; the protection of one's image and voice; and in being monitored and surveilled.[120]

In her comment on safeguarding human rights in the era of AI, Dunja Mijatović, the Council of Europe Commissioner for Human Rights puts it like this.

'The tension between advantages of AI technology and risks for our human rights becomes most evident in the field of privacy. Privacy is a fundamental human right, essential in order to live in dignity and security. But in the digital environment, including when we use apps and social media platforms, large amounts of personal data are collected – with or without our knowledge – and can be used to profile us, and produce predictions of our behaviours. We

provide data on our health, political ideas and family life without knowing who is going to use this data, for what purposes and how.'[121]

AI depends on the collection and processing of vast amounts of data which can potentially include personal and even sensitive data. Anonymisation techniques – which have served as a wildcard, allowing companies to process anonymised personal data – could be more easily circumvented with the use of AI. Only a very small amount of data is needed to uniquely identify an individual. It was found that 87% of the population in the United States could be uniquely identified based only on cross-referencing a 5-digit ZIP code, gender, and date of birth.[122] AI exponentially strengthens data processing capabilities. If anonymised personal data becomes part of a large data set, AI can de-anonymise this data based on inferences from cross-referencing information. This blurs the distinction between personal and non-personal data, which is a cornerstone of present legislation.[123]

AI can also identify and extrapolate trends and patterns of behaviour which can be used to influence opinions, choices, and decisions.[124] According to Liesl Yearsley, former CEO of Cognea, a company working in the fields of deep learning and AI, it is extremely easy to exert influence over people's decisions, even with relatively simple programs: 'Every behavioral change we at Cognea wanted, we got it,' and this can be largely exploited by companies looking for profit, she said.[125] In doing so, there is also the potential for data to be misused or for AI to malfunction, which could cause harm and/or interfere with the rights of data subjects.

It will be important to see how the application of the EU's General Data Protection Regulation (GDPR) will impact profiling practices, including with the use of AI, and on automated decision-making that impacts individuals. According to article 22 of the GDPR, '[t]he data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.'[126] Moreover, according to article 29 Data Protection Working Party, 'to qualify as human intervention, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture.'[127]

There are three exceptions to the general prohibition on automated decision established by the GDPR.

Article 22 will not apply when the decision is (a) provided by the law, such as in the case of fraud prevention or money laundering checks; (b) necessary for the performance of or entering into a contract; or (3) based on the individual's prior consent. Nevertheless, even in these cases, the data controller needs to inform the data subject about the existence of automated decision-making, providing meaningful information about the logic involved,[128] as well as the significance and the envisaged consequences of such processing for the data subject.[129]

As can be inferred from the current regulation, data protection principles such as consent, fairness, purpose limitation, and data minimisation remain important operational safeguards for the exercise and enjoyment of the right to privacy, but also need to be fine-tuned with regard to the introduction of AI technologies. An AI model which produces discriminatory results could fail the test of fairness. An AI model which is not frugal in the data it collects (because it has not made a lateral sweep of the amount and nature of the information needed) is not proportionate and could fail the test of data minimisation.[130] An AI model which gathers or reuses personal data from various sources yet does not inform, explain, and/or obtain the consent of the data subject, could fail the purpose limitation principle. As regards the principle of consent, a GDPR-compliant privacy information notice cannot adopt the form of a blank cheque covering any type of machine learning or AI technology. The information notice to the data subject needs to explain the main elements considered in reaching the decision, the source of the information obtained and their relevance.[131] In other words, these principles underscore the autonomy of and respect for the data subject by requiring, for example, explicit consent and clear proof of significant interference of the right.

Data protection is especially important for vulnerable groups whose ability and capacity to consent may be less developed and/or diminished. As stated by Rosa Kornfeld-Matte, the UN independent expert on the enjoyment of all human rights by older persons: 'Older persons should also be able to change their minds and opt out of technology at any time. Unless there are viable alternatives, the older person does not, however, have a real choice. Consent is not merely an administrative requirement. It is an essential element to a rights-based approach.'[132] She underlines the importance of the provision of simple and accurate information about the technology in order for them to be able to assess its implications.[133]

In summary, the right to privacy is of critical importance for individuals who are increasingly surrounded by AI-enabled products and services in public, at work, or at home. It is a frontier field for the exercise of other rights and freedoms, which almost conditions their enjoyment. It depends on the operationalisation of laws and principles which protect data attributable to the data subject used by AI models.

The prohibition of discrimination

The ICCPR states: '(A)ll persons are equal before the law and are entitled without any discrimination to the equal protection of the law. In this respect, the law shall prohibit any discrimination and guarantee to all persons equal and effective protection against discrimination on any ground such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.'[134] This is a general prohibition, embedded in international human rights law, which states are obliged to protect. It is also a protection which is associated with the exercise of a wide range of other human rights. In practice, discrimination and unequal treatment take different forms including those based on gender, sexual orientation, and disability.

At the European level, it is the established case law of the European Court of Human Rights that a difference in treatment is discriminatory when it has no objective and reasonable justification, i.e., if it does not pursue a legitimate aim or if there is not a reasonable relationship of proportionality between the means employed and the aim sought to be realised.[135] The decision to discriminate should be intentional or sufficiently explicit and the unequal treatment is not objectively or reasonably justified. While there can be lawful grounds for discrimination, the European Court of Human Rights is likely to consider certain grounds, such as race and ethnic origin, as being a particularly egregious kind of discrimination in need of special vigilance from public authorities and the courts.[136] In such cases, the state may have a positive obligation to adopt measures to prohibit discrimination, even when this may concern private parties, such as technology companies.

Irrespective of where it may be found, it is the *potential* of AI to discriminate which matters, even if there is no underlying intention to do so. The question to be asked is whether AI-enabled products and services engender practices which are disadvantageous to certain categories of the population and which are not objectively or reasonably justified. The absence of law and case law makes this difficult to ascertain, especially if it can be argued that AI models risk being inherently discriminate in the way they function (i.e., when identifying, ranking, prioritising, associating, and filtering data) and/or risk standardising bias from datasets. To this end, the UN High Commissioner for Human Rights highlighted 'that as the use of artificial intelligence systems becomes more pervasive, there may be disproportionate and disparate impacts on certain groups facing systemic inequalities, including women within those groups.'[137,138,139] An AI tool used by Amazon to review job applications, for example, did not rate candidates for software developer jobs and other technical posts in a gender-neutral way. Because Amazon's AI was trained to vet applicants by observing patterns in past applications – most submitted by men – the system taught itself that male candidates were preferable.[140] We know that

'[m]achines function on the basis of what humans tell them. If a system is fed with human biases (conscious or unconscious) the result will inevitably be biased. The lack of diversity and inclusion in the design of AI systems is therefore a key concern: instead of making our decisions more objective, they could reinforce discrimination and prejudices by giving them an appearance of objectivity. There is increasing evidence that women, ethnic minorities, people with disabilities and LGBTI persons particularly suffer from discrimination by biased algorithms.'[141]

In practice, AI-enabled products and services can determine, intentionally or otherwise, what information people see and how they are perceived. UN Women pointed this out in a campaign that raised awareness about sexist results, ranging from stereotyping as well as outright denial of women's rights, that were being generated by the Google search autocomplete function, which is powered by AI.[142] Research has revealed the failure of certain AI facial recognition technologies to unequally mis-identify people of different skin colours and genders because the datasets that were used to train the AI system encompassed a larger number of white male individuals than black people and women.[143] Biased datasets have also negatively influenced Microsoft's

AI chat bot 'Tay' designed to learn from engaging with Twitter users in casual conversations. As people starting tweeting the bot with racist and sexist remarks, Tay started repeating these sentiments back to users.[144] In a similar example, MIT's Media Lab 'Norman', an algorithm designed to analyse pictures, which was trained on data from 'the dark corners of the net' started to interpret abstract images consistently as dead bodies, blood, and destruction.[145]

Therefore, the quality of data used by AI models is crucial. AI systems, trained on data which replicates existing racial and gender stereotypes, tend to amplify and perpetuate discriminatory practices which can interfere with the exercise and enjoyment of other human rights. Consider also the propensity to collect, aggregate, de-anonymise, and repurpose data (with a loss of the original context for its collection)[146] from an increasing array of data from different objects in the home and on the body (such as personal home assistants and fitness trackers), from which inferences can be drawn.

According to a study published by researchers at the University of California – Berkeley, AI is capable of identifying individuals by learning daily patterns, such as the number of steps collected by activity trackers, smartwatches, and smartphones, and correlating it with demographic data.[147] This information could be sold to third parties, such as health insurance companies, which could potentially discriminate against less physically active individuals.

The impact that discrimination can have on other human rights, such as the right to work and the free choice of employment, and to a standard of living with adequate health and wellbeing is also significant. With more data available, AI models continue to learn (associate and predict), in particular to correlate, deduce, and predict for commercial advantage. The outcomes of these models are economically valuable in employment,[148] insurance, and credit decisions. Some companies take into account the social network connections of loan applicants or the information collected by the geolocation feature of their mobile phones before deciding on granting a loan.[149] The use of AI could enable the analysis and cross-referencing of information in a much wider scale, as the ongoing implementation China's social rating system may demonstrate. The system is based on a series of big data and AI-enabled processes that effectively grants subjects a social credit score based on their social, political, and economic behaviour.[150]

AI models being used to assist law enforcement in predictive policing and to guide the courts in sentencing terms can be argued to interfere with the fundamental right to liberty and security[151] as well as the right to a fair trial.[152] AI-enabled decision-making in matters which can affect individuals so acutely must be fair and correct, making any interference with their human rights so critically important.

On this basis, any state claiming a pioneer role in the development and/or reliance on new technologies bears a special responsibility for striking the right balance between rights, duties, and interests.[153] With greater unthinking reliance on autonomous everyday AI, it can be argued that they have a positive obligation to anticipate and prevent existing and new forms of discrimination.

The right to hold opinions and to freedom of expression

Article 19 of the ICCPR states that everyone has the right to hold opinions without interference as well as the right to freedom of expression, which includes the freedom to seek, receive, and impart information and ideas of all kinds, regardless of frontiers, orally, in writing, or in print, in the form of art, or through any other media of their choice. Further, the exercise of this right carries with it special duties and responsibilities which may be subject to certain restrictions, but only such as are provided by law and are necessary for respect of the rights or reputations of others, and for the protection of national security or of public order (*ordre public*), or of public health or morals.

As an essential basis of a democratic society,[154] this right has been developed by the courts over the years to promote pluralism and diversity in information and expression. It has also resulted *inter alia* in affording the media the freedom to act as a watchdog over the public interest, in protecting human rights defenders, and in circumscribing limits to expression with regard to hate speech. The positive and negative obligations of states have been instrumental in this regard, especially in striking the right balance between competing rights, freedoms, and responsibilities.

The digital age has been instrumental in enabling unprecedented levels of freedom of expression and access to information. This has underscored the importance of access to the Internet provided and facilitated by providers of telecommunications and digital platforms. This has resulted in widespread support for Internet freedom and a light touch to regulating intermediaries. AI models are deployed by such intermediaries to manage vast amounts of data and to effectively help

users seek and access information (e.g. the resolution of search queries).

However, big-data (predictive) analytics, the profiling of users, and the personalisation of content, as well as automated filtering mechanisms, make AI models powerful agents for interference with freedom of expression due to their capability to de-index (prioritise and promote), filter, and takedown content. AI models have competing priorities to rank user-generated content without censoring it, and to monitor and take down unlawful, illegal, or otherwise harmful content without over-blocking it.

Paradoxically, AI models are also misused to autonomously generate fake news (e.g. on social media platforms such as YouTube[155]) in order to spread malicious disinformation which may be considered to interfere with the right to form opinions necessary in a democratic society (e.g. voting in elections). In all cases of AI-enabled expression, be it deliberate, unintentional, or inadvertent yet opportunistic in nature, the potential to shape public opinion and electoral processes is clear for all to see.

Any interference with freedom of expression and access to information performed by AI models, however inadvertent or for legitimate reasons (e.g. to comply with law enforcement requests), must meet the tests of legality, proportionality, and necessity. The state is obliged to ensure that any interferences are prescribed by law which is clear and foreseeable, entails action which is proportionate (e.g. does not lead to overbroad restrictions), and meets a pressing social need.

Although we have yet to see the courts determinedly address AI interferences with freedom of expression, certain precedents can already be extrapolated from other judgments:

- AI should not result in the blanket blocking of access to online content.[156]
- AI-enabled search listings should not interfere with the reputation of others.[157]
- AI-enabled media have responsibilities to moderate comments they openly receive which may lead to hate speech and incitement to violence.[158]

AI-enabled freedom of expression necessitates, on the one hand, an awareness of its importance for individual liberty, society, and the public interest (i.e., the freedom to offend, shock, or disturb)[159] whilst, on the other hand, remaining sensitive to the need to protect individuals and groups faced with hate speech or speech considered to be hateful and/or which incites violence. Consequently, the state should be wary of affording AI-dependent companies the unfettered freedom to facilitate, moderate, and even influence expression and information, especially when this is performed by autonomous machines in a manner which is not transparent or explainable.

Conclusions

AI has entered a digital landscape which lacks clear regulation and guidance from case law. It will quietly learn and increasingly assist people for good in their everyday lives.[160] AI models will drive us around, help us make choices, and complement the care for loved ones. In this regard, they will enable the effective exercise and enjoyment of human rights for many more people than ever before. At the same time, they are likely to engender risks *inter alia* to life, to liberty, to dignity, and to work, many of which are not yet known nor fully understood.[161]

Considering the fast pace of technological change, it will be insufficient (too little, too late) to rely on traditional mechanisms and powers to regulate AI models as they arrive.[162] It is therefore incumbent on states relying on AI to maintain continuous reflection and deliberation on the human rights impact of AI, for example by means of open and vigorous public debate. There should also be mechanisms available to challenge and/or correct AI inputs and outputs regardless of their source[163] as well as capacity building efforts (e.g. online training courses on AI) and dialogue with companies on the design, development, and delivery of AI-enabled products and services.

Conclusions and further recommendations
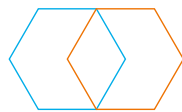
This report provides a broad overview of the relationship between artificial intelligence (AI) and diplomacy. It is clear that AI is becoming a prominent topic on the international agenda due to its broad potential impact on many aspects of our social, political, and economic lives. It is also clear that, in one form or another, AI impacts the diplomatic agenda and that it will transform a number of debates and introduce new, and sometimes very challenging topics. Ministries of foreign affairs (MFAs) have begun to prepare for the shifts caused by AI by releasing or beginning to work towards national AI strategies. Some MFAs have also begun to explore AI applications as tools for diplomatic practice.

In conclusion, we make three final recommendations for MFAs and diplomats to get ready to engage with the topics and tools emerging from AI's impact on international relations and the practice of diplomacy. We previously issued similar recommendations regarding capacity building and organisational adjustments in the context of big data. Not only are these recommendations also applicable when it comes to AI, they are in fact all the more important.

Develop indicators for tracking *diplomacy for AI*
- National AI strategies and related reports set or suggest goals for countries. However, clear and measurable objectives are missing in many cases. Hence, the next step for many countries will be to clearly define goals and operationalise them in order to develop targets and indicators.[164]
- There are at least three reasons why this will be important. First, given the substantial investment in the building of national AI ecosystems, data infrastructures, and AI applications, concrete indicators will contribute to accountability towards citizens. Second, operationalisation will help in tracking progress and making adjustments where necessary. Third, since many of the developments in AI are in the very early stages, testing (and revising) initial assumptions and documenting lessons learned will be crucial.[165]

- As national AI strategies get more refined, a clearer idea of the foreign policy goals related to AI will emerge (outcome indicators) and diplomats and MFAs will be tasked with working towards these goals. In this regard, it will be important to define what the MFA needs in order to effectively support the implementation of national AI strategies. The following questions can offer a first guidance in order to define input indicators for MFAs:
  – What is the number of existing diplomatic personnel with a background in topics relevant for AI?
  – What is the number of diplomatic personnel receiving topic-specific training?
  – Do dedicated personnel exist who can support or focus on negotiations on topics with relevance for AI – either associated with key topical areas or as an independent cross-cutting unit?

Engage in capacity building
- To get ready to address AI effectively, existing capacities and capacity gaps within MFAs need to be carefully assessed. As AI is a new topic that is sometimes difficult to grasp given overstated dystopian as well as utopian reporting, diplomats need a basic understanding that will help them navigate the topic.
- Carefully designed capacity building that aims at supporting the skills and knowledge of diplomats is important. A foundation level of knowledge that enables diplomats to critically assess opportunities and challenges associated with AI and to put various applications in context is paramount. When it comes to working with AI as a tool, those diplomats directly involved need a more detailed understanding and need to be able to communicate effectively with software developers who are working on an AI application.

Making organisational adjustments: Building a cross-cutting unit on AI applications
- As MFAs begin to explore AI as a tool for their own work, it will become important to determine how

this process can be organised and fit within the existing organisational structure of ministries. It is clear that the development of AI applications for use by diplomats can benefit a variety of departments and units, cutting across regional and thematic departments.

- Taking inspiration from the Open Source Unit of the UK Foreign and Commonwealth Office and its work in the area of big data, we suggest the creation of a small cross-cutting unit on AI applications. This should be a small and agile unit with a multidisciplinary team, whose members have expertise in computer science, social science, law, and other disciplines. The unit should not only be able to develop AI applications but also to critically assess their implications and feasibility.
- In support of this, we suggest the identification of AI champions across the ministry. These are diplomats who are knowledgeable of or interested in AI and can facilitate the conversation between the unit and the other departments in the ministry.

As many of the issues we discussed in this report are still very much under development, it is not easy to foresee where the debate on AI and diplomacy will lead in the coming years. As we said in the introduction, diplomacy is here to stay. But diplomacy will also have to adapt to new challenges.

Notes

1. Floridi L (2012) Hyperhistories and the philosophy of information policies. *Philosophy & Technology* 25(2), p. 129.

2. Pinker S (1995) *The Language Instinct. How the Mind Creates Language.* London, UK: Penguin, pp. 192–193.

3. Elkus A (2016) The futility of (narrow) speculation about machines and jobs. *Medium* blog, 19 October. Available at https://medium.com/@Aelkus/the-futility-of-narrow-speculation-about-machines-and-jobs-a116a0672659 [accessed 29 November 2018].

4. Future Today Institute (2018) *2018 Tech Trends Report.* Available at https://futuretodayinstitute.com/2018-tech-trends-annual-report/ [accessed 29 November 2018], p. 47

5. Skilton M and Hovsepian F (2018) *The Fourth Industrial Revolution. Responding to the Impact of Artificial Intelligence on Business.* Cham, Switzerland: Palgrave Macmillan, p. 80.

6. Marr B (2018) The key definitions of artificial intelligence (AI) that explain its importance. *Forbes*, 14 February. Available at https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#3b6263494f5d [accessed 29 November 2018].

7. McCarthy J, Minsky ML, Rochester N, and Shannon CE (1955) A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Available at http://people.csail.mit.edu/brooks/idocs/DartmouthProposal.pdf [accessed 7 December 2018].

8. While IBM Watson won against human players, it also made mistakes that a human player would not have made. This serves as a useful reminder that artificial and human intelligence are, in essence, not comparable. For this point see also Tutt A (2016) An FDA for Algorithms. *69 Administrative Law Review,* 83, pp. 84–123. Available at https://ssrn.com/abstract=2747994 [accessed 8 December 2018].

9. Neapolitan RE and Jian X (2018) *Artificial Intelligence. With an Introduction to Machine Learning*, 2nd ed, Boca Raton, FL: CRC Press, p. 3.

10. Kumar E (2008) *Artificial Intelligence.* New Delhi, India: I.K. International Publishing House, p. 14

11. New Scientist (2018) DeepMind's Go playing software can now beat you at two more games. Available at https://www.newscientist.com/article/2187599-deepminds-go-playing-software-can-now-beat-you-at-two-more-games/ [accessed 7 December 2018].

12. Cerf V (2016) Artificial Intelligence and the Internet. Open Transcripts. Available at http://opentranscripts.org/transcript/artificial-intelligence-and-internet/ [accessed 7 December 2018].

13. Walsh T (2018) *Machines that Think. The Future of Artificial Intelligence.* Amherst, NY: Prometheus.

14. The Whitehouse (2016) *The Administration's Report on the Future of Artificial Intelligence.* Available at https://obamawhitehouse.archives.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence [accessed 7 December 2018].

15. Della Carva M (2018, 3 January) Elon Musk says AI could doom human civilization. Zuckerberg disagrees. Who's right? *USA Today*. Available at https://eu.usatoday.com/story/tech/news/2018/01/02/artificial-intelligence-end-world-overblown-fears/985813001/ [accessed 7 December 2018].

16. Brooks R (2017) Machine learning explained. *MIT* blog, 28 August. Available at https://rodneybrooks.com/forai-machine-learning-explained/ [accessed 7 December 2018].

17. Walsh T (2018) *Machines that Think. The Future of Artificial Intelligence.* Amherst, NY: Prometheus.

18. For an overview of different kinds of algorithms see Rosen Jacobson B, Höne KE, and Kurbalija J (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era.* Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 30 November 2018], p. 18.

19. Grossfeld B (2017) A simple way to understand machine learning vs deep learning. *Zendesk* blog, 18 July. Available at https://www.zendesk.com/blog/machine-learning-and-deep-learning/ [accessed 30 November 2018].

20. Marcus G (2018) The deepest problem with deep learning. *Medium* blog, 1 December. Available at https://medium.com/@GaryMarcus/the-deepest-problem-with-deep-learning-91c5991f5695 [accessed 7 December 2018].

21. Lovy D (2018) AI made really, really simple. *Medium* blog, 2 February. Available at https://medium.com/@danlovy1961/ai-made-really-really-simple-698da599bc00 [accessed 7 December 2018].

22. For a detailed description on, for example, the behaviour of neurons within the network and the use of training algorithms see MIT Press (2018) How smart machines really think. *Medium* blog, 9 November. Available at https://medium.com/@mitpress/how-smart-machines-really-think-a2b4a67db248 [accessed 23 December 2018].

23. Silver D *et al.* (2017) Mastering the game of *Go* without human knowledge. *Nature* 550, pp. 354–359. Available at https://www.nature.com/articles/nature24270 [accessed 7 December 2018].

24. Sing S, Okun A, and Jackson A (2017) Learning to play Go from scratch. *Nature* 550, pp. 336–337. Available at https://www.nature.com/articles/550336a [accessed 7 December 2018].

25. Silver D *et al.* (2018) A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362(6419), pp. 1140–1144. Available at http://science.sciencemag.org/content/362/6419/1140 [accessed 7 December 2018].

26. Tutt A (2016) An FDA for Algorithms. *69 Administrative Law Review* 83, pp. 84–123. Available at https://ssrn.com/abstract=2747994 [accessed 8 December 2018].

27. Rosen Jacobson B, Höne KE, and Kurbalija J (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era.* Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 30 November 2018].

28. Walsh T (2018) *Machines that Think. The Future of Artificial Intelligence.* Amherst, NY: Prometheus.

29. ITU News (2018) Challenges and opportunities of artificial intelligence for good. Available at https://news.itu.int/challenges-and-opportunities-of-artificial-intelligence-for-good/ [accessed 29 November 2018].

30. For the following overview, compare Scott B, Heumann S, and Lorenz P (2018) Artificial intelligence and foreign policy. *Stiftung Neue Verantwortung Policy Brief.* Available at https://www.stiftung-nv.de/sites/default/files/ai_foreign_policy.pdf [accessed 6 December 2018].

31. Government of Sweden (2018) *AI in the Nordic-Baltic region.* Available at https://www.regeringen.se/49a602/globalassets/regeringen/dokument/naringsdepartementet/20180514_nmr_deklaration-slutlig-webb.pdf [accessed 7 December 2018].

32. United Nations (1961) *Vienna Convention on Diplomatic Relations*, article 3. Available at https://treaties.un.org/doc/Treaties/1964/06/19640624%2002-10%20AM/Ch_III_3p.pdf [accessed 7 December 2018].

33. Neumann IB (2008) Diplomacy and globalisation. In Cooper AF, Hocking B, and Maley W (eds) (2008) *Global Governance and Diplomacy. Worlds Apart?* Basingstoke, UK: Palgrave Macmillan, pp. 15–28.

34. Wight M and Butterfield H (1969). *Diplomatic Investigations. Essays in the Theory of International Politics.* London, UK: Allen & Unwin.

35. Kurbalija J (2016) From harmonising cyberpolicies to promoting twiplomacy: How diplomacy can strengthen Asia-Europe's digital connectivity. In Asia Europe Foundation (ed). *Connectivity: Facts and Perspectives,* Volume I, p. 63. Available at https://www.diplomacy.edu/sites/default/files/ASEF%20Outlook%20Report%20Kurbalija%20Chapter.pdf [accessed 27 December 2018].

36. Zwetsloot R, Toner H, and Ding J (2018) Beyond the AI arms race. *Foreign Affairs.* Available at https://www.foreignaffairs.com/reviews/review-essay/2018-11-16/beyond-ai-arms-race [accessed 27 December 2018].

37. Mazarr MJ *et al.* (2018) *Understanding the Emerging Era of International Competition: Theoretical and Historical Perspectives.* Santa Monica, CA: RAND Corporation. Available at https://www.rand.org/pubs/research_reports/RR2726.html [accessed 27 December 2018].

38. Kurbalija J (2017) The impact of (big) data on geopolitics, negotiations, and the diplomatic modus operandi. *Diplo* blog, 19 May.

Available at https://www.diplomacy.edu/blog/impact-big-da-ta-geopolitics-negotiations-and-diplomatic-modus-operandi [accessed 27 December 2018].

39 Copeland D (2016) Science Diplomacy, In Constantinou CM, Kerr P, and Sharp P (eds) *The Sage Handbook of Diplomacy.* Thousand Oaks, CA: Sage.

40 The Royal Society and the American Association for the Advancement of Science (2010) *New Frontiers in Science Diplomacy: Navigating the Changing Balance of Power,* p. vi. Available at https://royalsociety.org/~/media/royal_society_content/policy/publications/2010/4294969468.pdf [accessed 7 December 2018],

41 Copeland D (2016) Science Diplomacy, In Constantinou CM, Kerr P, and Sharp P (eds) *The Sage Handbook of Diplomacy.* Thousand Oaks, CA: Sage, p. 629.

42 Berkman PA (2018) Could science diplomacy be the key to stabilising international relations? *The Conversation*, 12 June. Available at https://theconversation.com/could-science-diplo-macy-be-the-key-to-stabilizing-international-relations-87836 [accessed 7 December 2018].

43 European Commission (2018) *EU Declaration on Cooperation on Artificial Intelligence.* Available at https://ec.europa.eu/jrc/communities/en/community/digitranscope-digital-transfor-mation-and-governance-human-society/document/eu-decla-ration [accessed 7 December 2018].

44 To create the overview contained in this chapter, national AI strat-egies, white papers, or reports from the following countries were analysed: Australia, Canada, China, Finland, France, Germany, India, Italy, Japan, Mexico, New Zealand, the joint declaration of the Nordic countries (Denmark, Estonia, Finland, the Faroe Islands, Iceland, Latvia, Lithuania, Norway, Sweden, and the Åland Islands), South Korea, Sweden, and the United Kingdom.

45 Boston Consulting Group (2018) *Mind the (AI) Gap. Leadership makes the Difference.* Available at http://image-src.bcg.com/Images/Mind_the%28AI%29Gap-Focus_tcm108-208965.pdf [accessed 7 December 2018].

46 Dutton T (2018) *An Overview of National AI Strategies.* Available at https://medium.com/politics-ai/an-overview-of-nation-al-ai-strategies-2a70ec6edfd [accessed 12 December 2018].

47 UK Department for Business, Energy, and Industrial Strategy and Department for Digital Culture Media and Sport (2018) *AI Sector Deal.* Available at https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal [accessed 12 December 2018].

48 Ministry of Enterprise and Innovation (2018) *National Approach to Artificial Intelligence,* p. 5. Available at https://www.regerin-gen.se/4aa638/contentassets/a6488ccebc6f418e9ad-a18bae40bb71f/national-approach-to-artificial-intelligence.pdf [accessed 12 December 2018].

49 AI Taiwan (no date) Cabinet plans to develop the nation's AI industry. Available at https://ai.taiwan.gov.tw/news/cabi-net-plans-to-develop-the-nations-ai-industry/ [accessed 12 December 2018].

50 German Federal Government (2018) *Artificial Intelligence Strategy,* p. 8. Available at https://www.ki-strategie-deutsch-land.de/home.html [accessed 12 December 2018].

51 Webster G *et al.* (2017) Full Translation: China's 'New Generation Artificial Intelligence Development Plan'. *New America* blog, 1 August. Available at https://www.newamerica.org/cybersecu-rity-initiative/digichina/blog/full-translation-chinas-new-gen-eration-artificial-intelligence-development-plan-2017/ [accessed 12 December 2018].

52 Heumann S (2018) *Nationale KI-Strategien in EU und USA: Wie posi-tioniert sich die KI-Enquete?* Available at https://www.stiftung-nv.de/sites/default/files/folien_heumann_ki_enquete.pdf [accessed 12 December 2018].

53 CIFAR (2017) *Pan-Canadian Artificial Intelligence Strategy.* Available at https://www.cifar.ca/ai/pan-canadian-artificial-in-telligence-strategy [accessed 7 December 2018].

54 New J (2018) *Why the United States needs a National Artificial Intelligence Strategy and What it Should Look Like.* Available at http://www2.datainnovation.org/2018-national-ai-strategy.pdf [accessed 12 December 2018].

55 Ibid.

56 Australian Government (2017) *Australia 2030. Prosperity through Innovation,* p. 61.Available at https://industry.gov.au/

Innovation-and-Science-Australia/Documents/Australia-2030-Prosperity-through-Innovation-Full-Report.pdf [accessed 12 December 2018].

57 Webster G *et al.* (2017) Full translation: China's 'New Generation Artificial Intelligence Development Plan.' *New America* blog, 1 August. Available at https://www.newamerica.org/cybersecu-rity-initiative/digichina/blog/full-translation-chinas-new-gen-eration-artificial-intelligence-development-plan-2017/ [accessed 12 December 2018].

58 Ministry of Economic Affairs and Employment (2017) *Finland's Age of Artificial Intelligence,* p. 19. Available at http://julkai-sut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAl-lowed=y [accessed 12 December 2018].

59 Ibid. p. 55.

60 Villani C (2017) *For a Meaningful Artificial Intelligence*, p. 8. Available at https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf [accessed 12 December 2018].

61 Ibid.

62 The Federal Government (2018) *Artificial Intelligence Strategy,* p. 33. Available at https://www.ki-strategie-deutschland.de/home.html [accessed 12 December 2018].

63 The National Institution for Transforming India (2018) National Strategy for Artificial Intelligence. *Discussion Paper,* p. 7. Available at https://www.niti.gov.in/writereaddata/files/doc-ument_publication/NationalStrategy-for-AI-Discussion-Paper.pdf [accessed 12 December 2018].

64 Agenzia per l'Italia Digitale (2018) *Making the Most of Artificial Intelligence in Italy and Europe.* Available at https://libro-bian-co-ia.readthedocs.io/en/latest/doc/making_the_most.html [accessed 12 December 2018].

65 Strategic Council for AI Strategy (2017) *Artificial Intelligence Technology Strategy,* p. 3. Available at https://www.nedo.go.jp/content/100865202.pdf [accessed 12 December 2018].

66 British Embassy Mexico City, Oxford Insights, C-Minds (2018) *Towards an AI Strategy in Mexico,* p. 8. Available at https://www.cminds.co/copy-of-ai [accessed 12 December 2018].

67 The AI Forum of New Zealand (2018) *Artificial Intelligence. Shaping a future New Zealand,* p. 84. Available at https://aiforum.org.nz/wp-content/uploads/2018/07/AI-Report-2018_web-version.pdf [accessed 12 December 2018].

68 Ibid, p. 84.

69 Government of Sweden (2018) *AI in the Nordic-Baltic Region,* p. 1. Available at https://www.regeringen.se/49a602/globalassets/regeringen/dokument/naringsdepartementet/20180514_nmr_deklaration-slutlig-webb.pdf [accessed 7 December 2018].

70 Government of the Republic of Korea (2017) *Mid- to Long-Term Master Plan for the Intelligent Information Society,* p. 34. Available at https://k-erc.eu/wp-content/uploads/2017/12/Master-Plan-for-the-intelligent-information-society.pdf [accessed 12 December 2018].

71 Government Offices of Sweden (2018) *National Approach to Artificial Intelligence,* p. 8. Available at https://www.regerin-gen.se/4aa638/contentassets/a6488ccebc6f418e9ad-a18bae40bb71f/national-approach-to-artificial-intelligence.pdf [accessed 12 December 2018].

72 House of Lords (2018) *AI in the UK: Ready, Willing and Able.* Available at https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf [accessed 12 December 2018].

73 HM Government (2018) *Industrial Strategy. Artificial Intelligence Sector Deal,* p. 8. Available at https://www.gov.uk/government/publications/artificial-intelligence-sector-deal [accessed 12 December 2018].

74 Ibid. p. 9..

75 The Federal Government (2018) *Artificial Intelligence Strategy,* p. 33. Available at https://www.ki-strategie-deutschland.de/home.html [accessed 12 December 2018].

76 Ibid., p. 38.

77 Government Offices of Sweden (2018) *National Approach to Artificial Intelligence,* p. 10. Available at https://www.regerin-gen.se/4aa638/contentassets/a6488ccebc6f418e9ad-a18bae40bb71f/national-approach-to-artificial-intelligence.pdf [accessed 12 December 2018].

78 Harhoff *et al.* (2018) Outline for a German Strategy for Artificial Intelligence. *Stiftung Neue Verantwortung Policy Brief.* https://

www.stiftung-nv.de/sites/default/files/outline_for_a_german_artificial_intelligence_strategy.pdf [accessed 4 December 2018].

79 Webster G *et al.* (2017) Full translation: China's 'New Generation Artificial Intelligence Development Plan.' *New America* blog, 1 August. Available at https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/ [accessed 12 December 2018].

80 Ministry of Economic Affairs and Employment (2017) *Finland's Age of Artificial Intelligence*, p. 23. Available at http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y [accessed 12 December 2018].

81 Villani C (2017) *For a meaningful Artificial Intelligence*, p. 49. Available at https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf [accessed 12 December 2018].

82 British Embassy Mexico City, Oxford Insights, C-Minds (2018) *Towards an AI Strategy in Mexico*, p. 32. Available at https://www.cminds.co/copy-of-ai [accessed 12 December 2018].

83 Government of Sweden (2018) *AI in the Nordic-Baltic Region*. Available at https://www.regeringen.se/49a602/globalassets/regeringen/dokument/naringsdepartementet/20180514_nmr_deklaration-slutlig-webb.pdf [accessed 7 December 2018].

84 Government Offices of Sweden (2018) *National Approach to Artificial Intelligence*, p. 7. Available at https://www.regeringen.se/4aa638/contentassets/a6488ccebc6f418e9ad-a18bae40bb71f/national-approach-to-artificial-intelligence.pdf [accessed 12 December 2018].

85 European Commission (2018) EU member states sign up to cooperate on artificial intelligence. *Digibyte* blog, 10 April. Available at https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence [accessed 12 December 2018].

86 European Commission (2018) The age of artificial intelligence. *EPSC Strategic Notes*, pp. 5 Available at https://ec.europa.eu/epsc/sites/epsc/files/epsc_strategicnote_ai.pdf [accessed 12 December 2018].

87 Ibid., pp. 8-9.

88 Ibid., p. 12.

89 Ibid., p. 6.

90 United Nations (2018) *UN Secretary-General's Strategy on New Technologies*. Available at http://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf [accessed 4 December 2018].

91 United Nations (2018) *Report of the High-level Committee on Programmes at its Thirty-Sixth Session* (CEB/2018/6). Available at https://www.unsceb.org/CEBPublicFiles/CEB-2018-6-HLCP36-SoC-Unedited%20Version.pdf [accessed 4 December 2018].

92 IGF (2018) IGF 2018 Speech by French President Emmanuel Macron. Available at http://www.intgovforum.org/multilingual/content/igf-2018-speech-by-french-president-emmanuel-macron [accessed 4 December 2018].

93 Chivot E (2018) French Proposal for an "IPCC of AI" will only set back global cooperation on AI. *Center for Data Innovation* blog, 11 December. Available at https://www.datainnovation.org/2018/12/french-proposal-for-an-ipcc-of-ai-will-only-set-back-global-cooperation-on-ai/ [accessed 20 December 2018].

94 Friend T (2018) How frightened should we be of A.I.? *The New Yorker*, 14 May. Available at https://www.newyorker.com/magazine/2018/05/14/how-frightened-should-we-be-of-ai [accessed 9 December 2018].

95 Bjola C (2017) Digital diplomacy 2.0 pushes the boundary. *Global Times*, 5 November. Available at http://www.globaltimes.cn/content/1073667.shtml?utm_content=buffer6475c&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer [accessed 9 December 2018].

96 Kurbalija J (1996) Information technology and diplomacy in a changing environment. *Diplomatic Studies Programme. Discussion Papers,* No 20, p. 37. Leicester, UK: Centre for the Study of Diplomacy.

97 Meserole C (2018) Artificial intelligence and the security dilemma. *Brookings* blog, 6 November. Available at https://www.brookings.edu/blog/order-from-chaos/2018/11/06/artificial-intelligence-and-the-security-dilemma/ [accessed 12 December 2018].

98 Wagner D and Furst K (2018) AI and the international relations of the future. *International Policy Digest* blog, 12 August. Available at https://intpolicydigest.org/2018/08/12/ai-and-the-international-relations-of-the-future/ [accessed 12 December 2018].

99 The exception is Dickow M and Jacob D (2018) The global debate on the future of artificial intelligence. *Stiftung Wissenschaft und Politik*. Available at https://www.swp-berlin.org/fileadmin/contents/products/comments/2018C23_dkw_job.pdf [accessed 12 December 2018].

100 PWC (2018) *Künstliche Intelligenz als Innovationsbeschleuniger in Unternehmen* [Artificial Intelligence as an Accelerator of Innovation in Business], p. 7. Available at https://www.pwc.co.uk/services/audit-assurance/risk-assurance/services/technology-risk/technology-risk-insights/accelerating-innovation-through-responsible-ai/changing-rules.html [accessed 9 December 2018].

101 Daugherty PR and Wilson HJ (2018) *Human + Machine: Reimagining Work in the Age of AI*. Boston, MA: Accenture Global Solutions [e-pub].

102 PWC (2018) *Künstliche Intelligenz als Innovationsbeschleuniger in Unternehmen* [Artificial Intelligence as an Accelerator of Innovation in Business], p. 7. Available at https://www.pwc.co.uk/services/audit-assurance/risk-assurance/services/technology-risk/technology-risk-insights/accelerating-innovation-through-responsible-ai/changing-rules.html [accessed 9 December 2018].

103 We introduced NLP in the first chapter. It stands for natural language processing. Applications include understanding and answering questions, translation between natural languages, and speech recognition.

104 Höne KE (2018) Shaping an intelligent tech & trade initiative – ITTI driving AI-trade interactions to boost global prosperity. *Digital Watch.* Available at https://dig.watch/resources/shaping-intelligent-tech-trade-initiative-itti-driving-ai-trade-interactions-boost [accessed 9 December 2018].

105 Aquino Santos MR (2018) Cognitive trading using Watson. *IBM* blog, 12 December. Available at https://www.ibm.com/blogs/bluemix/2018/12/cognitive-trading-using-watson/ [accessed 21 December 2018].

106 Höne KE (2018) Shaping an intelligent tech & trade initiative – ITTI driving AI-trade interactions to boost global prosperity. *Digital Watch.* Available at https://dig.watch/resources/shaping-intelligent-tech-trade-initiative-itti-driving-ai-trade-interactions-boost [accessed 9 December 2018].

107 Ibid.

108 Dickow M and Jacob D (2018) The global debate on the future of artificial intelligence. *Stiftung Wissenschaft und Politik*, p. 7. Available at https://www.swp-berlin.org/fileadmin/contents/products/comments/2018C23_dkw_job.pdf [accessed 12 December 2018].

109 Guterres A (2018) Address to the Paris Peace Forum. Available at https://www.un.org/sg/en/content/sg/speeches/2018-11-11/address-paris-peace-forum [accessed 9 December 2018].

110 Mijatović D (2018) In the era of artificial intelligence: safeguarding human rights. *Open Democracy* blog, 3 July. Available at https://www.opendemocracy.net/digitaliberties/dunja-mijatovi/in-era-of-artificial-intelligence-safeguarding-human-rights [accessed 9 December 2018].

111 World Summit on the Information Society (2005) *Tunis Agenda for the Information Society*. Available at http://www.itu.int/net/wsis/docs2/tunis/off/6rev1.html [accessed 9 December 2018].

112 United Nations Human Rights Council (2012) *The promotion, protection and enjoyment of human rights on the Internet* (A/HRC/20/L.13). Available at https://ap.ohchr.org/documents/E/HRC/d_res_dec/A_HRC_20_L13.doc [accessed 9 December 2018].

113 United Nations Human Rights Council (2014) *The Promotion, Protection, and Enjoyment of Human Rights on the Internet* (A/HRC/26/L.24). Available at http://ap.ohchr.org/documents/alldocs.aspx?doc_id=23480 [accessed 9 December 2018].

114 United Nations Human Rights Council (2016) *The Promotion, Protection, and Enjoyment of Human Rights on the Internet* (A/HRC/32/L.20). Available at http://undocs.org/A/HRC/32/L.2 [accessed 9 December 2018].

115 Kurbalija J (2018) A tipping point for the Internet: ten predictions for 2018. *DiploFoundation* blog, 11 January. Available at **https://www.diplomacy.edu/blog/2018predictions** [accessed 9 December 2018].

116 An illustrative example can be found in the case Big Brother Watch and Others v. the United Kingdom. The ruling from the European Court of Human Rights asserted that some surveillance programs under the UK's Investigatory Powers Act 2016 violated privacy rights enshrined in the European Convention on Human Rights. European Court of Human Rights (2018) *Case of Big Brother Watch and others v. the United Kingdom* (applications nos. 58170/13, 62322/14 and 24960/15). Available at **https://hudoc.echr.coe.int/eng#{"itemid":["001-186048"]}** [accessed 9 December 2018].

117 Kurbalija J (2018) A tipping point for the Internet: ten predictions for 2018. *DiploFoundation* blog, 11 January. Available at **https://www.diplomacy.edu/blog/2018predictions** [accessed 9 December 2018].

118 Council of Europe Guide (2014) *Guide to Human Rights for Internet Users*. Available at **https://www.coe.int/en/web/freedom-expression/guide-to-human-rights-for-internet-users** [accessed 9 December 2018].

119 OHCHR (1976) *International Covenant on Civil and Political Rights*. Available at **https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx** [accessed 9 December 2018].

120 For factsheets on case-law and pending cases see European Court of Human Rights (no date) *Factsheets*. Available at **https://www.echr.coe.int/Pages/home.aspx?p=press/factsheets&c** [accessed 9 December 2018].

121 Council of Europe (2018) *Safeguarding Human Rights in the Era of Artificial Intelligence.* Available at **https://www.coe.int/en/web/commissioner/-/safeguarding-human-rights-in-the-era-of-artificial-intelligence?inheritRedirect=true** [accessed 9 December 2018].

122 Golle P (2006) Revisiting the uniqueness of simple demographics in the US population. *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society*, pp. 77–80.

123 Lubarski B (2017) Re-identification of "anonymized" data. *Georgetown Law Technological Review*. Available at **https://georgetownlawtechreview.org/re-identification-of-anonymized-data/GLTR-04-2017/** [accessed 9 December 2018].

124 Levin S (2017) Face-reading AI will be able to detect your politics and IQ, professor says. *The Guardian*, 12 September. Available at **https://www.theguardian.com/technology/2017/sep/12/artificial-intelligence-face-recognition-michal-kosinski** [accessed 9 December 2018].

125 Yearsley L (2017) We need to talk about the power of AI to manipulate humans. *MIT Technology Review*, 5 June. Available at **https://www.technologyreview.com/s/608036/we-need-to-talk-about-the-power-of-ai-to-manipulate-humans/** [accessed 9 December 2018].

126 European Parliament and European Council (2016) *General Data Protection Regulation*. Art. 22. Available at **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679** [accessed 9 December 2018].

127 Article 29 Data Protection Working Party. *Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation* 2016/679. Adopted on 3 October 2017.

128 According to article 29 of the Data Protection Working party, the explanation of the logic involved would include details on the rationale behind, or the criteria relied on in reaching the decision, without necessarily always attempting a complex explanation of the algorithms used or disclosure of the full algorithm. Article 29 Data Protection Working Party. *Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation* 2016/679. Adopted on 3 October 2017.

129 European Parliament and European Council (2016) *General Data Protection Regulation*. Art. 13f. Available at **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679** [accessed 9 December 2018].

130 Norwegian Data Protection Authority (2017) *Artificial Intelligence and Privacy.* Available at **https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf** [accessed 9 December 2018].

131 Article 29 Data Protection Working Party. *Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation 2016/679.* Adopted on 3 October 2017, part D 'Rights of the data subject'.

132 United Nations Human Rights Council (2017) *Report of the Independent Expert on the Enjoyment of all Human Rights by Older Persons* (A/HRC/36/48). Available at **https://undocs.org/A/HRC/36/48** [accessed 9 December 2018].

133 Ibid.

134 OHCHR (1976) *International Covenant on Civil and Political Rights*, art. 26. Available at **https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx** [accessed 9 December 2018].

135 European Commission (2011) *The Prohibition of Discrimination under European Human Rights Law. Relevance for the EU Non-discrimination Directives – An Update.* Available at **https://www.ab.gov.tr/files/ardb/evt/1_avrupa_birligi/1_6_raporlar/1_3_diger/human_rights/prohibition_of_discrimination_under_european_human_rights_law.pdf** [accessed 9 December 2018].

136 Murray D and Fussey P (2018) Police are using big data to profile young people, putting them at risk of discrimination. *The Human Rights, Big Data and Technology Project* blog, 17 May. Available at **https://www.hrbdt.ac.uk/police-are-using-big-data-to-profile-young-people-putting-them-at-risk-of-discrimination/** [accessed 9 December 2018].

137 United Nations Human Rights Council (2017) *The Promotion, Protection and Enjoyment of Human Rights on the Internet* (A/HRC/35/9). Available at **https://www.ohchr.org/Documents/Issues/Women/WRGS/A_HRC_35_9_AEV.docx** [accessed 9 December 2018].

138 Barocas S and Selbst AD (2016) Big data's disparate impact. *California Law Review* 104(67), pp. 671–732. Available at **https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899** [accessed 9 December 2018].

139 Boyd D, Levy K, and Marwick A (2014) The networked nature of algorithmic discrimination. In Ganghadaran S, Eubanks V, and Barocas S (eds) *Data and Discrimination: Collected Essays.* Washington D.C.: Open Technology Institute, pp. 53–57.

140 Dastin J (2018) Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, 10 October. Available at **https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G** [accessed 9 December 2018].

141 Mijatović D (2018) In the era of artificial intelligence: safeguarding human rights. *Open Democracy* blog, 3 July. Available at **https://www.opendemocracy.net/digitaliberties/dunja-mijatovi/in-era-of-artificial-intelligence-safeguarding-human-rights** [accessed 9 December 2018].

142 UN Women (2013) UN Women ad series reveals widespread sexism. Available at **http://www.unwomen.org/en/news/stories/2013/10/women-should-ads** [accessed 9 December 2018]. Dority H-J (2018) Artificial Intelligence and human rights – Why does it matter? *Medium* blog, 4 June. Available at **https://medium.com/@amnestyusa/artificial-intelligence-and-human-rights-why-does-it-matter-63f3cb9e2663** [accessed 9 December 2018].

143 Lohr S (2018) Facial recognition is accurate, if you're a white guy*. The New York Times*, 9 February. Available at **https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html** [accessed 9 December 2018].

144 Vincent J (2016) Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. *The Verge*, 24 March. Available at **https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist** [accessed 9 December 2018].

145 Wakefield J (2018) Are you scared yet? Meet Norman, the psychopathic AI. *BBC,* 2 June. Available at **https://www.bbc.com/news/technology-44040008** [accessed 9 December 2018].

146 Council of Europe (2017) *Algorithms and Human Rights.* Strasbourg: Council of Europe, p. 19. Available at **https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5** [accessed 9 December 2018].

147 Hickey J (2018) Advancement of artificial intelligence opens health data privacy to attack. *Berkeley News*, 21 December. Available at **https://news.berkeley.edu/2018/12/21/advancement-of-artificial-intelligence-opens-health-data-privacy-to-attack/** [accessed 26 December 2018].

148 UN General Assembly (1948) *Universal Declaration of Human Rights* (217 [III] A). Paris, 23(1).

149 Wadell K (2016) How algorithms can bring down minorities' credit scores. *The Atlantic*, 2 December. Available at https://www.theatlantic.com/technology/archive/2016/12/how-algorithms-can-bring-down-minorities-credit-scores/509333/ [accessed 7 December 2018].

150 Munro K (2018) China's social credit system 'could interfere in other nations' sovereignty'. *The Guardian*, 28 June. Available at https://www.theguardian.com/world/2018/jun/28/chinas-social-credit-system-could-interfere-in-other-nations-sovereignty [accessed 7 December 2018].

151 OHCHR (1976) *International Covenant on Civil and Political Rights.* Art. 9. Available at https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx [accessed 9 December 2018].

152 Ibid., arts 14 and 15.

153 European Court of Human Rights (2008) *S and Marber v. the United Kingdom* (applications nos. 30562/04 and 30566/04). Available at http://www.bailii.org/eu/cases/ECHR/2008/1581.html [accessed 9 December 2018].

154 European Court of Human Rights (1976) *Handyside v. the United Kingdom* (application no. 54993/72) Available at https://hudoc.echr.coe.int/eng#{%22dmdocnumber%22:[%22695376%22],%22itemid%22:[%22001-57499%22]} [accessed 9 December 2018].

155 Albright J (2017) FakeTube: AI-generated news on YouTube. *Medium* blog, 17 January. Available at https://medium.com/@d1gi/faketube-ai-generated-news-on-youtube-233ad46849f9 [accessed 9 December 2018].

156 The violation of the right to freedom of expression was addressed in European Court of Human Rights (2012) *Ahmet Yıldırım v Turkey* (application no. 3111/10). Available at https://hudoc.echr.coe.int/eng-press#{%22itemid%22:[%22003-4191041-4965778%22]} [accessed 9 December 2018].

157 Court of Justice of the European Union (2014) *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62012CJ0131&from=EN [accessed 9 December 2018].

158 European Court of Human Rights (2015) *Delfi As vs Estonia* (application no. 64569/09). Available at https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-155105%22]} [accessed 9 December 2018].

159 Notably the 'freedom to offend, shock or disturb the State or any sector of the population (s)uch are the demands of that pluralism, tolerance and broadmindedness without which there is no "democratic society"'.
European Court of Human Rights (2018) *Factsheet: Hate speech*. Available at https://www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf [accessed 9 December 2018].

160 Cullen D (2018) Why artificial intelligence is already a human rights issue. *Oxford Human Rights Hub*, 31 January. Available at http://ohrh.law.ox.ac.uk/why-artificial-intelligence-is-already-a-human-rights-issue/ [accessed 9 December 2018].

161 Office of the High Commissioner for Human Rights (2017) *Statement on Visit to the USA, by Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights*.* Available at https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=22533 [accessed 9 December 2018].

162 Hoang LN (2018) Il faut nous préparer au possible avènement d'une intelligence artificielle de niveau humain. *Le Monde*, 17 October. Available at https://www.lemonde.fr/sciences/article/2018/10/17/il-faut-nous-preparer-au-possible-avenement-d-une-intelligence-artificielle-de-niveau-humain_5370653_1650684.html [accessed 9 December 2018].

163 Chander A (2017) *The Racist Algorithm.* Research paper No. 498, School of Law, University of California, Davis. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2795203 [accessed 9 December 2018].

164 The first paper to make suggestions along these lines is Heumann S and Zahn N (2018) Benchmarking National AI Strategies. Why and how indicators can support agile implementation. *Stiftung Neue Verantwortung Policy Brief.* Available at https://www.stiftung-nv.de/sites/default/files/benchmarking_ai_strategies.pdf [accessed 9 December 2018].

165 Crawford P and Bryce P (2003) Project monitoring and evaluation: a method for enhancing the efficiency and effectiveness of aid project implementation. *International Journal of Project Management* 21, pp. 363–373. Available at http://rozup.ir/up/paper/paper/1/a_method_for_enhancing_the_efficiency_and_effectiveness_of_aid_project_implementation.pdf [accessed 9 December 2018].