



Centre for International
Governance Innovation



WORKSHOP REPORT: AI, Global Governance and International Public Policy

Participants:

Ann Fitz-Gerald (BSIA), Odessa Primus (GARI), Michael Koran (GARI), Rohinton Medhora (CIGI), Andrew Thompson (BSIA), Scott Hamilton (BSIA), Jeremy Hunsinger (WLU), Thomas Ng (WLU), Jatin Nathwani (UW), Jesse Hoey (UW), Joel Blit (UW), Tamer Ozsu (UW), Branka Marijan (Project Ploughshares), Jessica West (Project Ploughshares), Kate Larson (UW), Fakhri Karray (UW), Bissan Ghaddar (Ivey), Jiri Sedivy (EDA), Tony Curzon-Price (UK Cabinet Office), Jason Lajoie (UW), Teodora Stirbat (GARI), Holger Hoos (Leiden University), Catherine Mulligan (UCL), Bob Fay (CIGI), Patrick van der Smagt (Volkswagen Group), Ratislav Kacer (Globesec), Aaron Shull (CIGI)

Introduction

This report outlines key takeaways and recommendations from the “Artificial Intelligence, Global Governance and International Public Policy Agenda” workshop that was hosted by the Balsillie School of International Affairs (BSIA), the Global Arena Research Institute (GARI) and the Centre for International Governance Innovation on 14 January, 2021, 9:00am – 1:00pm (EST).

Workshop Objectives

The workshop's main objectives were twofold: (1) To critically discuss policy and ethical challenges posed by Artificial Intelligence (AI)-driven technological developments; (2) To develop a broader research program supporting multi-disciplinary, transnational, and cutting-edge research, thought-leadership, and knowledge-mobilization in this growing field of study and practice. The workshop’s interdisciplinary discussion on AI will help to pave the way for future research and publications that integrate ‘STEM’ disciplines, topics, and scholars, with policy-making, natural-science, and social-science. The workshop's five topics and panel discussions were: (1) The Case for Human-Responsible AI Tech Development; (2) Democratization and AI;

(3) Digital Multilateralism; (4) The Great Divide; (5) Empowerment v.s. Fear and Control. These topics were selected based upon two previous workshops held in conjunction with the *Next100 symposium* series held by GARI in 2018 and 2019, as well as through discussions between organizing institutions, BSIA and GARI.

Attendees

The workshop involved approximately 20 leading scholars, senior policymakers, and industry leaders, all working at the intersection of AI and global governance. As this workshop is the first to connect and discuss AI and Global Governance with this unique cadre and assembly of minds, it is hoped that, in the future, participants will return and the list of experts and attendees will expand as workshops and meetings delving into these topics will help to facilitate novel discussions, dialogue, collaborations, networks, and results.

Introduction

In Western and industrialized states and societies, AI now impacts much of our lives in the form of digitalized economic, psychological, social, and political stimuli. In short: from our smartphones, to GPS navigation in cars, to our personal banking, communications, and leisure activities, we are now living in a digital world.

Two recent events have played an important role in highlighting the role of AI in our daily lives. First, the global COVID-19 pandemic has accelerated the establishment of a more virtual world, where schools, businesses, and even social and political relations, have been confined and relegated to online forms of communication and discussion through the digitized medium of social-media platforms such as Zoom and Facebook. Second, the USA's 46th Presidential Election in 2020, and digital media platform's role in pre-and post-election scenarios, highlighted the powerful and deep influence that digitized narratives and ideas have in shaping our 'physical' or material everyday reality. In this digitized online world, AI and other new digital technologies have significantly empowered 'big tech' firms such as Amazon, Facebook, Twitter, and Google, reaching new heights in both stock price and social influence. Yet, along with this unprecedented power and reach, an important implication of the rapid and deep spread of AI, big tech, and the

digitalization of the ‘everyday’ is a tangible feeling amongst ordinary citizens that big tech firms are assuming the government's role in regulating and structuring important aspects of daily life. Concerns that a new faith (and dependence) in ‘technology’ is replacing previous social placeholders of religion, family, and community, are commonplace.

The hybrid regime of AI, which exists at the global level, mainly consists of governments, big tech firms, multilateral institutions, and civil society. Several questions have been raised on the strategy each must engage to deal effectively with the governance of AI. Naturally, in Western liberal and democratic states, the public wants to know who, in the face of such technological power and digital spread, is actually making the real or impactful decisions among these bodies, and what exists as mechanisms of oversight and safeguarding so that no (digital or tech) power oversteps its bounds or threatens the security of citizens. For instance: What kind of AI technology is feasible in a properly-functioning democratic state, and what is not? What is the line between facilitating, or limiting, the agency necessary for democratic participation and individual freedom of thought? When does AI stop facilitating, and start hindering, democracy and dialogue?

The public is also interested in how to effectively and safely use AI, and to what ends it will be used for. What does the future hold for them in the spread of these technologies? Presently, there is a lack of understanding of how new innovations in digital world change the relationship between centres of political and economic power. For example, the invention, spread and increasing use of digital currency – ‘cryptocurrency’ – has provided an experiment showing us different ways to organize economic structures and institutions previously believed to have been limited solely to rates, treasuries, central banks of nation-states, and the businesses overlapping physical borders in a globalized web. However, these concerns about technology transforming sovereignty and economy, are not new; from the invention of the spinning jenny to the production line, scholars have long been deliberating how new technologies, such as AI, affect the economy and long-term (un)employment trends. Although the future is always unknown, we should learn lessons from our past to set progressive yet realistic policies for AI governance.

The Case for Human-Responsible AI/Tech Development

One of the key debates in the governance of all innovative technologies, but especially in AI, is what role humans should assume in the governance of these technologies. Innovation

benefits human beings, but when left unchecked, innovation can be counter-productive by making redundant or ineffective large swathes of technologies, workers, and social systems, that the continued operation of a state and society might rely upon for a peaceful and fluid operation. Policymakers, therefore, need to be cognizant that technological innovation is entwined in deep and complex ways with social and economic structures. Technologies like AI, machine learning, autonomous controls, etc., can, in many ways that appear to mirror a human's, think, predict, and act. In short, they can limit or replace a human beings' role in traditional decision-making processes. This raises the question of whether AI is more objective or transparent than a human decision-maker, or whether it (a) lacks social norms, ethics, and morals essential to integrating context and empathy into decisions, and (b) whether the algorithms operating its AI systems will merely replicate the underlying biases of its initial programmers. On the one hand, some scholars believe that the lead role of AI will contribute to neutrality, transparency, and efficiency, in social and political processes. On the other hand, however, critics argue that AI-based systems are also quite limited, and will indeed be biased in many subtle ways; manipulable, for instance, by humans selecting specific data sets or analysis systems as their foundation for action. Such a 'data dependency' in decision-making processes also eliminates qualitative or intangible elements of human psychology and selfhood, such as love, sympathy, friendship, and empathy within the governance system; qualities which are vital hallmarks of any civilized society. When a government applies a statistical model to a situation, it is using data from a specific time and place, yet it removes this context in order to generalize and expand its results upwards to be significant for other places, times, and peoples. Importantly, regardless of the data used and uncovered, this data must always be interpreted and 'thought' by a human being in a specific social and historical time and place. This is not to say, however, that statistics and AI cannot work in tandem with social systems, or cannot at least strive for objectivity. A case study from the United Kingdom (UK), for instance, can have lessons for governments in other parts of the world that seek to make better social policies through AI or algorithms. Due to the COVID-19 outbreak, many high school exams were canceled in the UK, and the Ministry of Education then allocated grades through a complex statistical procedure that allowed students to remain physically distance while still delivering grades. However, many people labeled this initiative as governance through AI, or overstepping the bounds, and unpredictability, of both educational testing and childhood development. A significant number of students protested and criticized the algorithmic procedure. Hence, these

types of intersections between social events and AI illustrate a stark division in societies about human reliance on technology and the role AI can play in altering traditional practices and patterns of behavior. Data integrity is also debatable; the system, for example, may intentionally be fed incorrect data to obtain a particular decision that might benefit one actor at the expense of another. Who is responsible for making a wrong decision in the presence of AI? If the human component is involved, accountability is easy, but if AI runs the system, it is not easy to fix many errors. For example, in autonomous cars, if accidents occur that result in damage, injury, and/or death, whom should therefore be held responsible for this accident? The designer, manufacturer, or programmer of the vehicle?

When we discuss ‘responsible AI’, we are now talking about transnational processes that affect peoples and states across the world. Yet the application of AI has different meanings for different societies, and hence the programming of decisions and even ethical behaviors, might differ. Policymakers must also be mindful that there is a plurality of ethical values and beliefs in the world (e.g. European AI might focus more on regulation and standards; American AI on individualism, rights, and freedom; Chinese AI on authority and adherence to group rules, etc.) and with these cultural attributes underlying the programming of algorithms, there may, therefore, be a ‘non-Western AI’ that emerges. Moreover, AI has a different impact from one technological and economic sector to another. In the defence sector, AI and machines will be making decisions about people's lives and deaths, which will make developers and end-users more conscious and fearful of how these innovative technologies make decisions and ascribe worth to living things. In other words, it is the "kill decisions" that make these things controversial. AI application in the education system or hiring processes will impact grades, or financial and health decisions, but will likely not determine specific or direct ‘life and death’ choices and outcomes. AI, therefore, deserves more critical attention when it comes to ethical perspectives and analyses; it can revolutionize modern warfare and impact new and traditional warfare techniques, but how it does so will ultimately hinge upon capacities for decision-making that are e-managed by human creators who retain their biases, histories, and modes of interpretation.

For example, the European Defence Agency (EDA) is responsible for supporting multilateral defence projects in the European Union (EU). At present, the EDA is carrying out 300 defence projects – and around 40 of these projects concern AI. A few examples of AI integration in EU's defence projects include: communication, intelligence, surveillance,

reconnaissance, management of command and control, decision-making, autonomous transport, and logistic systems. The EU has called for common international legal frameworks concerning AI, robotics, and autonomous systems. However, due to differences in legal traditions and legal cultures, there is a difference of opinion within EU member states. Within the EU, Germany has emerged as a leader for ethical scrutiny of innovative technologies like AI. A debate has emerged within the EU that reflects a common public desire to ensure that 'big tech' firms adequately address ethical and legal concerns, ranging from privacy rights, to the ownership of data, utilizing insights from technical experts and social scientists alike.

When it comes to warfare, many countries believe that maintaining human control of autonomous weapons systems is necessary. This debate is currently occurring in the UN and has been ongoing for a significant time, but a primary question sowing debate is: what type of control constitutes 'control', and what type of freedom constitutes autonomy? Some countries believe that, if humans are involved in the "kill decision," it is an acceptable practice to grant this power to AI; others argue that there should be more meaningful and direct human involvement in any kill decision, since morality and ethics are essential components of human thought and action. With this in mind, the question then becomes: what is a 'meaningful involvement' of humans in an AI system?

The EU is embracing the idea that, to integrate ethics within innovative technologies, it is better to move away from two-stage processes where technology is introduced, which enables social impact analysis to be done. There is a growing consensus that engineers should work closely with social scientists from the beginning of the project to its end. Multi-disciplinary collaboration may ameliorate the previous tendency for AI to lack, or circumvent, typical ethical processes. For data protection, the EU has already set up certain standards through General Data Protection Regulation (GDPR) and intends to enact more data governance and AI legislation going forth. As well, European Commission's ethics guidelines regarding AI are the basis of upcoming legislation regarding AI. All these measures indicate that Europe can lead the world in data protection and AI governance if it continues this trend of attention and allocation of resources. The technical investment has social payouts and ramifications; many experts from the EU believe that AI can be regulated through traditional international law and human rights frameworks despite being a new technology.

Recently, the Canadian government has announced increased rules and regulations around AI, and Canadian officials are keen to work in this area. Moreover, Canada has excellent laws on protecting health data. Unless explicit given consent from a person is obtained, their data needs to be, and is, protected. However, there are instances where third parties that are not regulated by the government have given health data to insurance companies, and that is where the government's role in protecting citizens and their privacy rights is crucial. An area where Canada needs improvement is indeed in the intersection of ethics and science. Not many computer science programs in Canada's universities are combining ethics and science to confront the complex decisions and developments outlined above, and hence, there is a need to reform the education system to take account of these ongoing and important technological and social shifts.

Democratization and AI

Democratization can be defined in many ways, but for AI, one possible application is that every citizen receives the opportunity and benefit of this resource or technology. In simplest terms, the democratization of AI means that AI is available to all people, for their use and benefit if they choose. Characteristics of democratization processes can include the following: First, AI's democratization means making it more accessible, which means designing AI to be easy to use by other domain experts. Institutions such as NGOs, governments, and Small and Medium Enterprises (SMEs), should not be far behind the big tech firms in knowledge and capacity related to AI. Second, there should be a consistent attempt to balance private financial and commercial entities' legitimate interests with public interests concerning AI. One way to understand the importance of AI in the commercial world is that it is not just a technology, but also a money-making tool for commercial entities. Unfortunately, it is also this possibility that could be a key hindrance in making AI more democratic, as entities gaining power and wealth through the benefits of this technology may attempt to withhold it from others.

Third, although an application of AI could perhaps be justified to attain efficiency, the legitimacy of its usage in determining democratic and political outcomes is critical for making the case that AI-related political processes are real and effective. Without this legitimacy, the use, spread, and governance of AI, will always be challengeable and suspect. Historically, liberal democracies generally deal with innovative technologies through 'exit and choice' options for the public; however, these options may not be readily available in the case of AI. Gaining and

maintaining legitimacy is a complex process, and it may have different meanings, for example, in the global North and in the global South. One way to obtain legitimacy of AI-based systems is to obtain stakeholders' ex-ante consent instead of ex-post-facto approval. Ex-post-facto change in the rules or procedures removes the citizenry's political agency, and so publics feel angry when they perceive that the rules of their governance system are changed to their detriment, without their consent. To some extent, the deployment of innovative technology should be similar to crucial rulemaking processes, in which stakeholders are consulted before passing an act or statute. For AI-based systems, the rules should be defined and consulted by many, before they come into force on a large-scale to address the stakeholder and societal concerns about distributional consequences. Moreover, the legal system's appeal and review process suggest that the AI-based systems should also be designed with certain review processes and remedial measures. In Canada, the Federal government has issued a directive on automated decision-making and citizens' right to explainability. For example, if a person is denied a benefit where an automated system was used as part of the judgment, the affected citizen has a right to seek explainability from the authorities.

Critics believe that AI is not fully democratized, and that only a few big tech firms and powerful countries have a monopoly over the use and design of AI. This makes the differences between accessibility, democratization, and ethics, essential, and their explanations and usages should be clear so as to reduce disagreement and debate. One probable reason for this debate is that existing models of capitalism and liberal democracy prefer short-term gains, making it challenging for policymakers to develop long-term policies that appear to cost more time, money, energy, and patience from the public, in the short-term. There is a need to find a mechanism that can guide and evaluate AI's long-term impact on our lives. One possible approach is to provide resources to universities to research and implement projects that analyze AI's long-term application for human beings. However, as noted above, most educational systems are not well equipped to help students understand both the ethics and technicalities of AI, and this field is still in its infancy. Researchers should study, for instance, whether AI is similar to many other innovations and technologies with which human beings have dealt with previously (e.g. the internal combustion engine; the internet; etc.), or if it has more potential for transformative impact on human life. Another way to democratize AI governance is to encourage the establishment of unions and protections for data governance and the ethical evaluation of AI. Organized unions, for instance,

might provide platforms for workers to raise their voice if there is a breach of trust from a big tech firm, a specific technology, or a governmental institution.

Digital multilateralism

What does multilateralism mean in the digital era? Technologies such as AI have global implications, but there is no global dialogue or platform for AI governance. The UN high-level report on digital cooperation is an excellent document to understand digital multilateralism. There is a broader understanding among the stakeholders that there is no unified view about digital multilateralism or multilateralism in the digital age. Although big tech firms have a somewhat better or more nuanced understanding of these complex issues, they are also working because of operating on the basis of self- and shareholder interest, rather than the public good. As a result of this, there is a global demand for multilateral rules and standard settings for AI and data governance, ensuring the public good and private industry can coexist. The fair distribution of economic gains through advancements and uses of AI can only be achieved if there exists an effective multilateral regulatory system to guide it and punish transgressions. In response to these demands, we have seen some initiatives like Global Partnership on AI (GPAI) 's creation in the Organisation for Economic Co-operation and Development (OECD). However, most of these partnerships are not global, but regional or national. Hence, there is a need for a truly global platform to address multiple issues at the national and international level simultaneously. Multilateralism on AI is essential because different regions use and develop different forms and types of data, ultimately resulting in different technologies and governance models.

There are three very different notions on how data can be used in digital governance. In the USA, data governance is private sector-focused and supports individual freedom, while in the EU, the focus is on ethics and strategic regulations through institutions like GDPR. In China's case, we see a great firewall that focuses on localization or widespread social benefits. Here, data brokers can easily sell data, because there are no stringent data privacy regulations.

An important question that matters in almost all multilateral issues is the relevance and effectiveness of the UN. Is the UN the best institution for multilateral governance of AI and digital cooperation? There are indicators that the UN is very active in this area, demonstrated, for instance, by the launch of its roadmap for digital cooperation. However, there are concerns that the UN may

not be ‘fit for the purpose’, as it lacks human resources with a strong interdisciplinary background dedicated to this emerging and rapidly developing field. Another challenge for the multilateral governance of AI through the UN is the desire of ‘great powers’ such as China and the USA, to be a norm or standard-setter for every new change, becoming leaders in the world through their specific governance model. Many argue that these countries do not necessarily want to see something develop successfully if it does not benefit their own strategic interests; rather, they prefer to shape structures for their interests, in what is known as a ‘zero-sum’ game. Moreover, countries with nationalistic agendas, like China, India, or the UK, can also resist calls for AI's multilateral governance by simply refusing to participate. However, the countries' positions change as their interests change and, in the case of AI, countries' positions will be evolving and the success of any governance platform will be to find the necessary and effective incentive that can allow this to happen. Although there is a need to differentiate between economic and non-economic issues in AI governance, there may be much more coherence on economic issues than the latter. One possible approach is finding the relatively ‘low hanging fruits’ for multilateral cooperation in areas such as policing the dark web, taxation regimes, reducing drug smuggling, terrorism, and finance.

Moreover, there are forums like the G-20 and the World Trade Organisation (WTO), which also offer a blueprint from which a multilateral governance framework for AI could be developed. The creation of the Financial Stability Board (FSB) by the G20 after the financial crisis of 2008 indicates that countries are indeed willing to cooperate and create new institutions to serve their interests. For example, in Waterloo, Canada, CIGI has adopted this initiative by establishing a Digital Stability Board (DSB) - a new organization that will bring together stakeholders to coordinate standards, regulations, and policies regarding AI and digital cooperation. The Board will monitor developments, advise on best practices, assess civil society's impact, determine the impact of these innovative technologies, provide policy and regulatory advice, and include voices that need to be heard. The key purpose of DSB is to avert a digital or AI crisis as we had in the financial sector in 2007-2008. There is a need to overcome national and societal differences so as to create a space of like-minded nations to share best practice and regulations, and establish ethical and behavioral norms, so that the rest of the world may follow in a stable and secure manner.

The Great Divide

The ‘great divide’ is generally defined as the gulf between haves and have-nots, encompassing the lack of access to AI's knowledge, technology, capacity, and infrastructure between various stakeholders. It is a multifaceted concept, and its implications are different from one sector to another. The expanding AI regime is a classic case of ‘insider and outsider’, in which those outside the regime do not have the capacity or voice in to influence decision-making processes, while those ‘inside’ are afforded power through the capability to determine the direction of AI’s development and spread. The great divide has revealed issues concerning trust, transparency, and accountability between users and service providers of this technology. It piggybacks on the common ‘digital divide’ which exists between countries and within countries (i.e. some citizens have access to technologies while other do not). As a result, the divide has exacerbated the power and reach of monopoly ‘big tech’ firms located within but a few developed countries. The emergence of a digital monopoly of data that is derived from global sources, but processed in a handful of countries, highlights the inequitable sharing of data, and hence, of knowledge and control. There is, therefore, a concern that existing foundations and developments around AI are based on systemic discrimination because of these divides. Colonialism, racism, and misogyny are the historical bedrock upon which many of the ‘big tech’ companies have staked their national operations, and without addressing these underlying legacies, there is a danger that AI outputs will only strengthen these subtle systemic biases.

A digital divide exists between big tech firms, small firms, and governments. As data and technology available to Amazon or Google are not available to small firms or governments, there is an unlevel playing field that is disadvantageous to citizens. As customers, citizens and users of a technology expect good service from firms such Google or Amazon. Conflict with nation-states arise because these big tech firms have set higher service delivery standards than the governments that regulate their reach, since the latter do not have the same resources and capacity to provide similar services. As a result, citizens get frustrated with governmental provisions, and agree to give more power and rights to big tech firms. Discrepancies of power also arise between nations as well. The lack of capacity of developing countries becomes a major hindrance in their effective negotiations with developed countries or big tech firms to secure their legitimate interests in digital sectors.

There is no mechanism to assess how anti-competitive behavior about data governance can be managed at the international level because all competitive laws exist at the national level. The role of governments, civil society, academia, and NGOs is important to push the big tech firms to disclose and publicize their data. Government institutions need to increase their capacity to deal with AI's complex issues from multiple aspects that include the histories of racism, colonialism, and misogyny, noted above. For example, the UK's competition commission criticized Google's decision to ban cross-domain cookies, while the information commission appreciated it. This kind of situation, where governmental institutions lack coherence in their policies and foster confusion rather than confidence and control, actually helps big tech firms to enforce their policies instead of following the government's regulatory guidelines.

Another aspect of the digital divide is the ability of AI to eliminate the jobs and reduce the workforce of a large number of people in the coming years. Many scholars disagree that AI will enormously destroy worker's employment, as they argue that, before the beginning of the Information Technology (IT) revolution, similar concerns were also shared, but not realized. Now, however, IT is one of the biggest employment sectors, and the actors, stakes, and reach of technology, have changed. Governments should therefore devise policies that ensure that AI's costs and benefits are evenly distributed among skilled and non-skilled workers.

On the other hand, proponents of AI believe that ICT has now bridged the digital divide, as more software and hardware tools such as faster computers and information clouds (i.e. storage devices) are available to the public. Due to this accessibility, there exists the potential for citizens to develop AI-based products in their homes and offices. This capacity for everyday persons to learn and use a technology is something different from the limited-access of the past, and credit must be given to big tech firms that have ensured the accessibility of these technologies to the public. For example, at the University of Waterloo (UW), math and engineering faculties have collaborated to develop an AI institute in partnership with business organizations to provide AI-based services to these business entities. It is a win-win situation for all the stakeholders.

Empowerment vs. Fear Control

Like many other modern technologies, AI has the potential to empower human beings and to hurt them. For example, AI will be enormously beneficial for disabled people and people with a critical

illness that require extra assistance, while able-bodied persons could possibly enhance their physical and decision-making capabilities. Some experts believe that AI is the most significant development since the nuclear age. However, there is a major governance difference between AI and nuclear technology, because governments have little control over the development, marketization, and spread of AI, unlike that of nuclear technology which remains (largely) determined by the state due to its cost and the sheer size and space required for a functional production facility. Moreover, AI and the digital world have empowered citizens in terms of the capacity for communication, and while sitting in remote or isolated areas of a country, they can engage social media, 'tweet', and convey their concerns to powerful elite such as the Prime Minister or president of a country. Indeed, it is in these latter areas that proponents of AI recommend its use and development (e.g. A vaccine-rollout conducted and calculated by AI may perhaps be more effective, and conclusive, than different layers of national governments pitting decision-makers against one another).

However, this enormous connectivity and empowerment due to AI comes with certain concerns and fears about privacy and security. If decisions are being made by big tech firms that impact our day-to-day lives without any regulatory framework, concerns over intrusion, overreach, and the selling of one's own private data to other companies and agents, naturally arises. We must acknowledge that fear is a basic human instinct and quite relevant to the emergence of AI, which is a new and largely unseen technology. If social media platforms such as Twitter, Facebook, or WhatsApp provide a sense of accessibility and empowerment, they also create a fear of surveillance. This fear arises because we do not know much about the AI entities monitoring and compiling this data, and this uncertainty or 'fear factor' is then multiplied when we read news reports stating that autocratic regimes are using spy software to track their critics and dissidents.

The implications of AI in the security arena are very significant. It has completely changed the intelligence world. Earlier intelligence was an exclusive domain of governments, and intelligence agencies were accountable to citizenry, to some extent, in democratic governments. However, big tech firms are now involved networks of surveillance and tracking, altering not only the means and mediums through which surveillance occurs, but also to whom results are reported and information is delivered. ICTs and data science are essential tools of modern warfare, and the allegations of foreign involvement in processes such as the USA's presidential electoral processes

indicate both the intensity and the nature of modern warfare: not only has it become digital, but it is even removed from human control. The lack of accountability of big tech firms in this form of espionage or warfare also makes the situation more complicated, as the discussion above on ethics and responsibility, indicates. However, the supporters of AI advancement argue that AI is not the only opaque component of our daily lives. Political systems and decision-making processes are, and have always, been complex and lacking in transparency.

One solution to manage the fear factor regarding AI is to establish a regulated and standard-setting process that applies across states. The most challenging part of regulating AI, however, is what must be regulated, where, and who actually has the power to enforce these regulations. AI governance does not necessarily mean developing a new governance infrastructure separate from political and economic structures, and in some sectors existing governance options may also work well in tandem or partnership. For example, the USA has mostly adequate governance infrastructure for managing the rollout of autonomous vehicles, yet things will clearly be different for systems of autonomous weapons. One way to develop effective regulatory regimes is that governments must become more active in this expanding domain, so that big tech firms are not the only actors to attract the Industry's best minds. For this purpose, the governments will have to allocate more budget resources for research and development concerning AI. Public entities should have the same technological and governance capabilities as the private sector to hold private entities accountable in AI applications. The recent COVID-19 pandemic has provided an historic opportunity to governments, allowing them to reimagine their role in a post-pandemic world, and to redefine the social contract that exists with their populace.

The following steps may be helpful for the guidance of policymakers:

Recommendations for Policy Makers:

- (1) Develop a multidisciplinary framework to deal with AI governance and move away from 'disciplinary silos' or conversations limited to one specific discipline or field. There is a need to develop interdisciplinary law(s) to regulate innovative technologies like AI.
- (2) Policymakers need to invest significant resources and time in close collaboration with the private sector to comprehend AI's costs and benefits for their respective countries.

- (3) Policymakers must ensure that national AI frameworks are consistent with universal governance frameworks, such as human rights declarations and environmental treaties and agreements. This can be difficult because strong states will typically advance their specific agendas in multilateral arenas, threatening the viability, or quieting the voices, of smaller or less-powerful states.
- (4) One possible approach for developing an ethical framework for AI may be John Rawls's concept of "reflective equilibrium." Reflective equilibrium is a process of revising our judgments according to debate, dialogue, and circumstance, achieved amongst a variety of different agents and actors.
- (5) The education system is a key resource to ensure a balanced relationship between AI and human beings. There is a need to revive the curriculum of specialized streams and make them more interdisciplinary. For example, medical students should be familiar with AI's basic concepts and their implications in their area of expertise. STEM, social science, and arts and humanities, should have a better-integrated voice on topics of AI.
- (6) More participation of technical experts in discussions (e.g. such as this workshop) will assist technical experts to understand technology's social implications. Ethics should be part of technology from the design stage to implementation.
- (7) Civil society should be empowered to act as a watchdog to ensure responsible innovation of AI.
- (8) No single framework will be helpful, but a multi-governance approach will be required for AI governance.
- (9) Each country's needs and ethical considerations will be different concerning the application of AI to their populace. Therefore, big tech firms and multilateral institutes should be cognizant of this diversity, and ready to adjust to it.
- (10) AI's application in modern warfare and autonomous weapons development may have irreversible adverse consequences and, as such, needs careful evaluation.

Future Research Program

- (1) Is international law adequate to deal with AI?
- (2) How will AI change the traditional power structures in International Relations?

- (3) In the presence of AI, how will governments work with the private sector?
- (4) How can AI help us understand and deal with complex global questions like climate change, poverty, and inequality?
- (5) Is our education system fit for coping with AI-related developments?
- (6) How can citizens have meaningful representation in the decision-making processes related to AI?
- (7) What are the capacity gaps between the global South and the global North regarding AI?
How can these gaps be filled?
- (8) What value can a transatlantic group on AI bring to the multilateral governance of AI?
- (9) What is the difference between consent and informed consent regarding AI? What are the implications for ordinary citizens?