ASSER INSTITUTE
Centre for International & European Law

WINTER ACADEMY ON ARTIFICIAL INTELLIGENCE AND INTERNATIONAL LAW

FEBRUARY 2019

REPORT

WINTER ACADEMY ON ARTIFICIAL INTELLIGENCE AND INTERNATIONAL LAW (2019)

REPORT

The first edition of the Asser Institute's Winter Academy on Artificial Intelligence and International Law took place in February 2019 in The Hague. It gathered over 25 participants from around the world, who received insights from, and engaged with, 20 expert speakers. The following Report contains brief summaries of each of the session.[1] The detailed programme of the 2019 Winter Academy is provided in Annex.

MONDAY 11 FEBRUARY 2019 – UNDERSTANDING AI

Session 1: History of Artificial Intelligence, Current Trends and Prospective Trajectories

Dr Giovanni Sileno (University of Amsterdam)

The Winter Academy started with a lecture by Dr Giovanni Sileno, who began by defining the terms 'artificial' and 'intelligence'. Dr Sileno then outlined and examined the four most popular approaches to the study of artificial intelligence (AI): the Turing test approach; the cognitive modelling approach; the laws of though approach; and the rational agent approach.

Dr Sileno mapped the 'start' of artificial intelligence, from the Dartmouth Conference in 1956 and the group of academics who coined the term, to the other factors that propelled its growth, such as the operational research in the United Kingdom (UK) in the 1930s and the transformation in psychology from the focus on explanation, instead of behaviourism.

He then went on to explain the meaning of computational intelligence, as well as the different styles of programming. Further, he illustrated the current status of AI research today, and also speculated on its potential trajectory, which he believed included cognitive architectures and 'artificially dumber' systems as (potentially) the next wave of AI.

Dr Sileno concluded by pointing out that human beings have always evolved being shaped by their tools, and therefore we should look at our tools (such as AI), as being not only means, but as forces that impact on and shape our societies.

Session 2: Individuality and Humanity in the Age of AI

Dr Haroon Sheikh (Wetenschappelijke Raad voor het Regeringsbeleid)

Dr Sheikh's lecture took a philosophical-speculative approach to the study of artificial intelligence. He speculated on how AI could impact and affect our worldview, and the elements of modernity that could be challenged by the rise of AI.

Dr Sheikh outlined the two prevailing views of how humans relate to technology. The 'technology as a tool' approach, which was put forward by René Descartes, views technology as a tool to provide use with safety, comfort, empowerment and freedom. The second

[1] This Report was drafted by Kike Ajibade, intern at the Asser Institute and participant to the 2019 Winter Academy, and Berenice Boutin, researcher at the Asser Institute and organiser of the 2019 Winter Academy. Any error or omission is borne by the Report's authors.

approach – 'technology as a trick', instead views technology as a destructive force, associated with artificiality and loss of contact with the world and ourselves. This view of technology, and AI is prevalent in popular culture (i.e. The Terminator, Black Mirror, etc.), but also has ancient origins (i.e. Frankenstein and Faust). Dr Sheikh explained that the underlying issue with both of these approaches is that they view humans and technology as separate, when we should instead view technology as something that is inherently involved within human nature.

According to Dr Sheikh, there were three elements that we associated with the modern age that could be undermined or transformed by the impact of new technologies such as AI: our idea of ourselves as autonomous individuals; our reliance on rational justification; and the view that mankind is the active agent in a passive world – 'human exceptionalism'.

In his final reflections, he noted that we were moving away from a world of individuality, and towards one of interconnectedness; away from a world of rationality and towards one of superstition, and finally away from a world in which humans are the exceptional beings, and towards one in which we are just one of many types of agents. He reiterated the need to see AI, and technology in general, as connected to us, and acknowledge the need to develop and evolve together with technology.

Session 3: Opacity and Discrimination in Algorithmic Decision-Making

Professor Frederik Zuiderveen Borgesius (Radboud University Nijmegen; University of Amsterdam)

Professor Borgesius' lecture addressed the issue of opacity in many AI systems, and the possibility of unfair discrimination related to this issue. He highlighted a few examples of such discrimination in everyday life. For example, one study carried out found that searches containing African-American names have a much higher chance of being shown ads related to arrest records than searches containing typically 'white-sounding' names. Another study by Datta et al using fake Gmail accounts found that female Internet users received fewer advertisements for higher-paying job compared to their male counterparts.

Professor Borgesius noted that the 'black box' nature of most machine learning systems made any investigation into discrimination very difficult, as it is impossible to ascertain where exactly the discriminatory effects were coming from. He explained that AI systems merely reproduced patterns that already existed in society.

Turning to the role of the law in alleviating these discriminatory effects, Professor Borgesius outlined the two most relevant fields of law in this area: data protection law (more specifically the GDPR and its rules concerning automated decision-making), and non-discrimination law (both direct and non-direct discrimination). He concluded by providing some recommendations for the future concerning how to mitigate the discriminatory effects of current AI systems.

Session 4: Hands-on Workshop on AI

Dr Giovanni Sileno (University of Amsterdam)

Dr Sileno's second presentation of the day provided a practical beginner's guide to artificial intelligence, coding, and how autonomous systems are actually made. Using easily accessible online resources, and recognisable, easy-to-understand examples, such as the Simpson's family tree, he taught the participants the fundamentals of coding. He also provided resources where they could learn more about coding and practice their skills if they wished.

TUESDAY 12 FEBRUARY 2019 – AI FOR GOOD

Session 5: Ethics of AI

Professor Jeroen van den Hoven (TU Delft)

Professor van den Hoven's presentation focused on the core problems related to the ethics of artificial intelligence. His lecture sought to map the contours of an emerging ethical framework in the field of AI.

He emphasised the need to move the discussions about AI away from the question of 'can computers think', and instead redirect the focus to the consequences of this powerful tool, and the extensive effects it can have in the real world.

Professor van den Hoven emphasised that the *designing* of the world was becoming more important, compared to the focus on explanation and prediction in the past. He gave examples of the political implications present in many artefacts, such as low-hanging overpasses, hostile architecture, and Jeremy Bentham's famous Panopticon prison design. In today's world, these methods of implementing ideas through design can be seen in the 'architecture of the internet'.

He ended by highlighting the effects that AI could have on our privacy, freedom, autonomy and democracy. He used the example of 'nudging', a concept in behavioural psychology which becomes more effective and powerful when combined with big data and high processing power.

Session 6: Human Rights and AI

Sherif Elsayed-Ali (Element AI)

Sherif Elsayed-Ali began by emphasising the fact that artificial intelligence was merely a specialised tool, which could be used for good or evil. AI cannot, in itself, violate human rights or cause global issues.

He outlined six key human rights issues that were linked to AI use: privacy; bias and discrimination; propaganda, fake news and manipulation; autonomous weapon systems; job losses and inequality.

He further explained ways in which AI could be used to actually protect and advance human rights and improve global welfare. He gave the example of a 2018 study conducted with Amnesty Tech, which used crowdsourcing and data science to analyse online abuse against women in the public sphere. Mr Elsayed-Ali noted that, although this system was merely a demo, such systems would eventually become good enough to assist human moderators in filtering out abusive messages on social media platforms.

Session 7: Putting Human Rights at the Heart of the Design, Development and Deployment of AI

Vivian Ng (University of Essex)

Vivian Ng's presentation addressed two questions: how does artificial intelligence affect our human rights, and how does the human rights framework respond to the challenges of AI and connect with an ethics-based approach of AI.

Ms Ng observed that the impact on privacy was one of the most well-documented effects of AI in society. However, she stressed that privacy had a gatekeeping function to many other

areas of our lives, and therefore the effects of AI could stretch to more fundamental rights, and affect our behaviour within society.

Ms Ng explained that the human rights framework was an important base point for considering these issues because these are rights that we enjoy inherently, simply by virtue of being human. Further, both substantive and procedural rights are included in the human rights framework, which covers a whole spectrum of rights, from civil and political to economic and social. Human rights contain tools and mechanisms for accountability that we can utilise to regulate these emerging technologies.

In the second part of her lecture, Ms Ng drew parallels between human rights and ethics-based approaches to AI, highlighting their commonalities, such as the focus on dignity and transparency. What separated these two approaches was the common language of harm added by the human rights approach, which gives us a more comprehensive and effective approach to identifying, defining and assessing harms done by AI.

Session 8: Securing Evidence of War Crimes with AI and Blockchain
Dearbhla Minogue (GLAN Law)

Using the example of the conflict underway in Yemen since 2015, this lecture by Dearbhla Minogue explored the use of artificial intelligence and blockchain technology in securing evidence of war crimes.

After defining and explaining the requisite elements of a war crime, Ms Minogue highlighted the main issues faced by national and international courts in obtaining evidence of war crimes, and thus why war crimes prosecutions are so rare. She explained that the reliance on problematic and unreliable witness evidence made creating a case very difficult, due to factors such as poor human memory, PTSD and fear of retribution for testifying in court. Therefore, while witness evidence is essential in the judicial process, it needs to be supplemented by other kinds of evidence.

Using the example of the Bosnian camps in the 1990s, she outlined how video evidence could successfully be used to form a criminal case, and mitigate the issue of the lack of access investigators have to a crime scene during or immediately after the event. Ms Minogue also highlighted a case before the ICC, in which the arrest warrant was issued in part on the basis of videos posted online of the defendant carrying out executions.

The speaker further described the role that the principle of chain of custody plays in the use of digital evidence, and how her own organisation and their partners, GLAN Law and Syrian Archive, uses 'hashing' and blockchain to preserve this chain of custody and ensure that the evidence can be used in a criminal trial. She ended the lecture by briefly explaining the potential applications of machine learning in searching and filtering large archives of digital evidence.

Session 9: Hackathon Rewind: Using AI to Predict Land-Grabbing
Moses Emuze (Monkey Code Team; CGI Nederland)

The final presentation of the day was delivered by Moses Emuze, a member of the Monkey Code Team, who were the winners of the Hackathon for Good held in The Hague in November 2018. The Hackathon for Good brought together 27 teams of more than 22 nationalities to work on the development of innovative data solutions focused on humanitarian disasters, fake news, evidence, emergency funds and land grabbing.

Mr Emuze described how the Monkey Code Team developed an open source solution to prevent land grabbing, which was based on the combination of various data including from social media. He further discussed the logistics behind a hackathon and the process involved in bringing a team together and developing and presenting an idea.

WEDNESDAY 13 FEBRUARY 2019 – AI AND ARMED CONFLICT

Session 10: Autonomous Weapons Systems and Human Dignity
Dr Ozlem Ulgen (Birmingham City University)

Dr Ulgen's lecture addressed the ethical issues and considerations involved with the use of autonomous weapons systems (AWS) for the concept of human dignity. She explained that such systems were characterised by the use of AI and robotics in order to achieve varying levels of autonomy in their critical functions. Due to this autonomy, these weapon systems were increasingly removing a degree of human involvement in the decision-making process to use lethal force.

The presentation was organised into three sections: (1) understanding the philosophical and ethical underpinnings of the concept of human dignity; (2) identifying human dignity as involving two components — status and treatment of human beings, and (3) considering the impact of autonomous weapon systems on human dignity.

During the first section of the lecture, Dr Ulgen observed that the idea of an innate human worthiness existed in many different belief systems – from ethical, philosophical and religious. While human dignity has been conceptualised and given content under the human rights framework, its origins lie much deeper.

The second portion of the lecture drew on Immanuel Kant's moral theory on ethical conduct, which provides a rationale for rules based on human dignity as status and as respectful treatment of human beings. During the final portion of the lecture, Dr Ulgen observed how these notions are undermined with the development of AWS, as they create a hierarchy of human dignity and the use of these weapons often manifest as 'disgraceful punishment' to the victims. Dr Ulgen stressed that there was indeed a moral basis found in the Marten's clause preventing the assumption that an act or weapon not explicitly prohibited in law is therefore permissible.

Session 11: Autonomous Weapons Systems and International Law
Professor Mary Ellen O'Connell (University of Notre Dame)

Professor O'Connell's lecture addressed the implications of the use of lethal autonomous weapon systems (also known as LAWS) on our moral and legal understanding of the restrictions on killing under international humanitarian law (IHL), and international law more generally.

She began with an exploration of the historical background of the prohibition of war, beginning with pacifism in early Christianity, to St Augustine's just war theory, and finally Hugo Grotius' move away from natural law to a system of positive law. Positivists began to codify these principles that had been passed down from natural law scholars concerning rules on how to wage war and which weapons were permissible. Following the outbreak of the First World War, the Kellogg-Briand Pact was signed, and soon after the UN Charter brought the most widely recognised prohibition on the use of force in the form of Article 2(4). These principles

then paved the way for the 1980 Convention on Certain Conventional Weapons, whose purpose was to review weapons that fell under Article 36 of Additional Protocol I.

Professor O'Connell went on to discuss the rise of Realism as the dominant political ideology in the 20th century, and how this increased the focus on military power and advantage at the expense of morality and ethics. She concluded by arguing that the ever-increasing lax approach to the fundamental IHL principles of proportionality, distinction and humanity have led to an increasing acceptance of the possibility to use LAWS in warfare, and that the imposition of a ban on LAWS would help to re-establish moral and legal principles in this area.

Session 12: Knowing and Seeing the Combatant: Visuality and Targeting in International Law

Dr Amin Parsa (Lund University)

In this lecture, Dr Parsa highlighted some of the difficulties involved in targeting in warfare, especially in relation to the requirements of knowledge and the requirements of visibility, both of which are affected by the involvement of AI technologies. He addressed the major questions associated with the study of visuality in armed conflict, namely: How do we see? Who has the right to see? Which interpretation of images and videos are prioritised, and on the basis of what knowledge?

Dr Parsa began by explaining the fundamental principle of distinction, which requires that only combatants can be lawfully targeted during armed conflicts, while civilians must be protected. He highlighted the main provisions of the Geneva Conventions and their Additional Protocols which provided definitions and descriptions of combatants and civilians. Dr Parsa indicated that the dependency of the law of armed conflict on visual signifiers of a combatant or civilian was problematic, as, due to emerging technologies, the process of targeting looks very different today – it is no longer based on identifying individuals by their uniform, but has more to do with their chains of relationships and other factors.

In analysing how emerging technologies have been used to ascertain combatants in modern armed conflicts, he distinguished between the technologies related to seeing, and those related to knowing the combatant. He provided examples of technologies of knowing in the context of counterinsurgencies, such as SCIPR, NORA and RTRG. In terms of technologies of seeing, the most famous application of technology is real-time surveillance using drones.

He concluded by explaining the concept of 'visuality', which was the hierarchies of knowledge, modes of attention and prevailing assumptions which underpinned contemporary targeting practices.

Session 13: Panel Discussion: Human Control over Autonomous Military Technologies

Lt Col Bart van den Bosch (Netherlands Defense Academy; University of Amsterdam), Professor, Mary Ellen O'Connell (University of Notre Dame), Dr Ozlem Ulgen (Birmingham City University), Major Ilse Verdiesen (TU Delft; Netherland Armed Forces), moderated by Dr Sofia Stolk (Asser Institute; VU Amsterdam)

The final session of the day was a panel discussion on the topic of human control over autonomous military technologies that was opened to external participants. The panel began by each speaker presenting their views on the topic, followed by a discussion and a round of questions and answers.

Professor O'Connell took the view that, for the principle of humanity to truly be realised, this required a true human conscience, which could only be exercised by a human being. She argued

that Noel Sharkey's definition of human control is the one that should be followed when considering autonomous military technologies. This approach requires that the human commander should, inter alia, have active cognitive participation in the attack, have full contextual and situational awareness of the target area at the time of the attack, and have a means for the rapid suspension or abortion of the attack.

Dr Ulgen stressed the need for human control to be present at every stage and spectrum of the life cycle of autonomous weapons — both pre-deployment and post-deployment. She argued that the qualifier of 'meaningful' in the human control test would only be useful in practice if we focused on what the human operator is doing in order to keep control over the machine. She emphasised that the role of humans in targeting decisions, especially those concerning life or death, is very difficult to replace and subsidise with an AI system.

Major Verdiesen provided a brief overview of her research, which attempts to combine an engineer's perspective of autonomous weapons with a practitioner's perspective. She expressed doubts over the concept of human control, as it may involve an element of influencing the perspectives and outcomes of technology, however, she agreed for the need to assign responsibility in the event of unintended consequences flowing from the use of autonomous weapons.

Lt Col van den Bosch argued that, if an AI system could perform objectively 'better' than a human being at a certain task, then it could be preferable to delegate that task solely to the machine. He stressed that, often, human involvement in a task leads to unpredictability and mistakes, while machines are less likely to make such mistakes. Further, in the context of warfare, there are many judgments to make that are simply too complex and fast for a human operator to make effectively as compared to machines.

THURSDAY 14 FEBRUARY 2019 – AI AND RESPONSIBILITY

Session 14: The Challenge of Accountable AI
Dr Machiko Kanetake (Utrecht University)

For the first session of the day, Dr Kanetake discussed the challenges faced in attempting to regulate lethal autonomous weapon systems, and AI technologies in general. She approached the issue from many different angles, such as practical and philosophical.

She began the session by opening the floor to questions, asking the participants if they were in favour of the creation of a new international treaty or other regulatory framework to govern AI technologies, taking in particular the example of autonomous weapons and, if so, what this regulatory framework would look like. She then went on to explore the various ideological perspectives concerning the role of international law in regulating new technologies. For example, while Realists would view international law as an instrument of control, Liberalists would likely view it as a constraint on the agency of international actors.

Dr Kanetake also explored the views of Constructivist and Feminist thought on the issue of the regulation of autonomous weapons. According to constructivists, our role as international lawyers and researchers shapes and conditions our behaviour, and heavily informs our views when it comes to the regulation of emerging technologies. The feminist angle, on the other hand, emphasises how women are often disproportionately affected by these lethal autonomous weapons, and stresses the importance of bringing a fresh, and often under-

discussed perspective to the debate on autonomous weapons to avoid the conversation becoming polarised.

Session 15: AI and Individual Criminal Responsibility

Dr Marta Bo (Asser Institute; IHEID) and Abhimanyu George Jain (IHEID)

In this lecture, Dr Bo and Mr Jain explored how to assign individual criminal responsibility for breaches of IHL committed using lethal autonomous weapons. The first portion of the lecture explored the sources of IHL rules on the conduct of hostilities, and the differences between 'Hague law', which covers the protection of civilians in armed conflict, and 'Geneva law', which regulates the actual conduct of hostilities. Mr Jain especially examined the IHL rules on distinction and proportionality.

The second part of the lecture provided an exploration of the war crimes related to IHL rules on the conduct of hostilities. More specifically, she focused on grave breaches and breaches of international criminal law under the Rome Statute. In analysing the differing *mens rea* and *actus reus* requirements under the Rome Statue and Additional Protocol I, Dr Bo demonstrated how the comparatively higher *mens rea* requirement under the Rome Statute causes difficulties when discussing criminal responsibility for the use of fully autonomous weapon systems.

The final portion of the lecture considered more deeply the challenges of individual criminal responsibility for crimes committed with AI technologies. More specifically, Dr Bo and Mr Jain explored the problem of the alleged 'accountability gap' in international criminal law, which is caused by the difficulty in attributing responsibility for a crime committed with an autonomous weapon. They also explored the modes of responsibility that would be most appropriate in this context. Their proposed solutions included the concept of perpetration by means and command responsibility.

Session 16: AI and State Responsibility

Dr Berenice Boutin (Asser Institute)

Dr Boutin's lecture addressed the issue of the circumstances in which a state can incur responsibility in relation to violations of international law involving AI. The first section of the lecture addressed the issue of the responsibility of states for their own development and use of AI technology, while the second section addressed a state's responsibility in relation to the conduct of private actors. Dr Boutin emphasised that the focus on state responsibility did not undermine the study of individual criminal responsibility, as these responsibilities are concurrent.

Dr Boutin began by setting out the legal framework of state responsibility, as found in the *Chorzów Factory* case, and codified in the ARSIWA. As stated in Article 2 ARSIWA, the two factors necessary for the commission of an internationally wrongful act of a state are: (1) the act or omission must be attributable to that state, and (2) the act or omission must constitute a breach of an international obligation of that state.

The first part of the lecture was split into two sections, the first involving a discussion of the difficulties involved in the attribution of wrongful conduct involving the use of AI, and the second exploring the obligations of diligence in the development of AI technologies.

During the second part of the lecture, Dr Boutin turned to the role that private companies in the development of AI and the responsibility for breaches resulting from this technology. Dr Boutin explained that, while a state is not directly responsible for the conduct of private actors

(unless it has directed the actors), a state can be held indirectly responsible if it fails to ensure that these actors within its jurisdiction respect international norms.

Session 17: The Morality of Artificial Agents
Professor Massimo Durante (University of Turin)

Professor Durante's lecture took a multi-disciplinary approach to the study of artificial intelligence and the morality of artificial agents. He began by illustrating the importance of constructing and framing the issue properly, as answers are only given in the perimeter of the question. Therefore, when asking questions such as 'how do we attribute responsibility' or 'what is a moral agent', we must first ascertain the purpose for which we are posing the question.

Professor Durante highlighted three conceptual standpoints related to this topic: action, delegation and epistemology. When explaining action, he noted that action was no longer the exclusive prerogative of human beings, and AI was also no longer defined by intelligence, but by agency. In describing the second conceptual standpoint of delegation, Professor Durante highlighted the increasing level of delegation of tasks and decisions to machine in today's world. He made links between this increasing delegation and our 'forms of life', which has been informed by delegation and the rise of big tech companies. When describing the final conceptual standpoint, epistemology, the speaker emphasised the turn from 'explanation, and then building' to 'building, and then explaining'.

Session 18: Group Debate: Towards a Framework of Shared Responsibility for and of AI
Moderated by Dr Berenice Boutin (Asser Institute)

The final session of the day took the form of an interactive group debate amongst participants, moderated by Dr Boutin. Some of the questions for discussion included: Are there responsibility gaps in the current framework governing AI technologies? Do we need new rules to govern this area, or do we merely need to reinterpret and apply existing rules? Will international standards and institutions be strong enough to orientate this development? How do we avoid a diffusion of responsibility and ensure actual access to remedies for victims of autonomous weapon systems? During the debate, the issue of the granting of legal personality to advanced AI systems was also raised.

FRIDAY 15 FEBRUARY 2019 – AI GOVERNANCE

Session 19: The Geopolitics of Artificial Intelligence
Dr Tim Sweijs (The Hague Centre for Strategic Studies (HCSS))

Dr Sweijs's presentation provided an overview of the geopolitical impact of the increasing use artificial intelligence. Beginning with a definition of geopolitics and a background of the geopolitical patterns in the world, he observed that there has been a shifting center of economic power, beginning in Europe, and then shifting to the US, then to Russia, and heading for China in 2025. This leads to international competition, which is directly reflected in the world of big tech and artificial intelligence.

Using the examples of China, the US, Russia and Europe, Dr Sweijs provided an analytic framework of the AI programmes of these nations on the basis of scale, speed, coherent strategy, degree of centralisation and knowledge. He illustrated how these AI strategies differed, and highlighted the strengths and weaknesses of each one.

In the final section of his presentation, Dr Sweijs analysed the geopolitical impact that AI could have on our world and international order. He analysed factors such as the economic impact, socio-political impact and security impact of the increasing use of AI in today's world.

Dr Sweijs concluded the lecture by posing some questions for the participants to consider. These included: How can we mitigate the negative effects of AI on the labour market? What does an AI-driven liberal democratic model look like? How should big tech companies be regulated, and do governments have the expertise to regulate them?

Session 20: AI and Political Theory
Dr Haye Hazenberg (TU Delft)

Dr Hazenberg's lecture addressed the implications of artificial intelligence on our political and legal institutions. He argued that there was evidence that we are moving towards a 'digital lifeworld' (in the words of Jürgen Habermas), characterised by increasingly capable systems, increasingly integrated technology and an increasingly quantified society.

In exploring the digital 'upgrading' of political and legal institutions, Dr Hazenberg illustrated that data and code had been transformed into power, thus affecting the concept of the use of force in a society, the level of governmental scrutiny in a society, and the level of perception control.

Dr Hazenberg discussed the potential need to adapt the political systems to respond to and work within the context of our newly digitalised society. He illustrated a few examples of how such a 'digital democracy' would function, for example, a digital direct democracy, a digital data democracy, a digital wiki democracy or a digital AI democracy.

He concluded the presentation by challenging the idea of implementing a self-enforcing and adaptive law, due to the fact that history is unjust, and prejudices would merely be reflected in algorithmic injustices.

Session 21: The Role of International Norms and Institutions in the Governance of AI
Irakli Beridze (UNICRI Centre for AI and Robotics)

Irakli Berizde delivered the final lecture of the Winter Academy, and discussed the important role of international institutions and norms in governing the development and use of artificial intelligence technology. Mr Beridze began by mapping the hugely accelerated development of AI in a number of different sectors, from healthcare to the military. This development was largely pushed forward by large private sector investments, although engagement by states has become much more commonplace. Mr Beridze noted that the UN and other multinational organisations were comparatively slow in catching up to the technological advancements.

Moving on to the major risks involved with the use of AI, Mr Beridze discussed the increased occurrence of AI-enhanced crime, the use of lethal autonomous weapon systems, the job losses associated with AI, and the existential risk of 'superintelligence'.

In terms of the governance of AI, the burden currently falls mainly on non-governmental groups (such as The Future Society and the IEEE), and international organisations like the UN, as states can be hesitant in placing limits and regulations on AI research, for fear of stifling innovation and causing themselves to fall behind their foreign competitors. While a variety of panels and

expert groups have been established at the UN-level to discuss the impact of AI, concrete decision-making has yet to take place.

Mr Beridze went on to analyse the main challenges involved in the regulation of AI, which include the discrete nature of the technology, and the fact that it is also very diffused and geographically widespread. He concluded by questioning whether the current system of regulation, which was created in the post-WWII period and in vastly different circumstances, was indeed out-dated, or could be adapted to the modern context.

ANNEX: PROGRAMME

MONDAY 11 FEBRUARY 2019 – UNDERSTANDING AI

9:30 - 10:30	**Opening and Introductions** *Professor Janne Nijman (Asser Institute), Dr Berenice Boutin (Asser Institute)*
10:30 - 11:00	Coffee/Tea Break
11:00 - 12:00	**History of AI, Current Trends, Prospective Trajectories** *Dr Giovanni Sileno (University of Amsterdam)*
12:00 - 13:00	Lunch
13:00 - 14:00	**Individuality and Humanity in the Age of AI** *Dr Haroon Sheikh (Wetenschappelijke Raad voor het Regeringsbeleid)*
14:00 - 14:30	Coffee/Tea Break
14:30 - 15:30	**Opacity and Discrimination in Algorithmic Decision-Making** *Professor Frederik Zuiderveen Borgesius (Radboud University Nijmegen; University of Amsterdam)*
15:30 - 16:00	Coffee/Tea Break
16:00 - 17:00	**Hands-on Workshop on AI** *Dr Giovanni Sileno (University of Amsterdam)*
18:00 - 19:30	Welcome drinks Grand Cafe Victoria (Prins Hendrikplein 10, The Hague)

TUESDAY 12 FEBRUARY 2019 – AI FOR GOOD

9:30 - 10:30	**Ethics of AI** *Professor Jeroen van den Hoven (TU Delft)*
10:30 - 11:00	Coffee/Tea Break
11:00 - 12:00	**Human Rights and AI** *Sherif Elsayed-Ali (Element AI)*
12:00 - 13:00	Lunch
13:00 - 14:00	**Putting Human Rights at the Heart of the Design, Development and Deployment of AI** *Vivian Ng (University of Essex)*
14:00 - 14:30	Coffee/Tea Break
14:30 - 15:30	**Securing Evidence of War Crimes with AI and Blockchain** *Dearbhla Minogue (GLAN Law)*
15:30 - 16:00	Coffee/Tea Break
16:00 - 17:00	**Hackathon Rewind: Using AI to Predict Land-Grabbing** *Monkey Code Team (CGI)*

WEDNESDAY 13 FEBRUARY 2019 – AI AND ARMED CONFLICT

9:30 - 10:30 **Autonomous Weapons Systems and Human Dignity**
Dr Ozlem Ulgen (Birmingham City University)

10:30 - 11:00 Coffee/Tea Break

11:00 - 12:00 **Autonomous Weapons Systems and International Humanitarian Law**
Professor Mary Ellen O'Connell (University of Notre Dame)

12:00 - 13:00 Lunch

13:00 - 14:00 **Knowing and Seeing the Combatant: Visuality and Targeting in International Law**
Dr Amin Parsa (Lund University)

14:00 - 14:30 Coffee/Tea Break

14:30 - 17:00 **Panel Discussion: Human Control over Autonomous Military Technologies**
- *Lt Col Bart van den Bosch (Netherlands Defence Academy; University of Amsterdam)*
- *Professor Mary Ellen O'Connell (University of Notre Dame)*
- *Dr Ozlem Ulgen (Birmingham City University)*
- *Major Ilse Verdiesen (TU Delft, Netherlands Armed Forces)*
- *Moderator: Dr Sofia Stolk (Asser Institute; VU Amsterdam)*

17:00 - 18:30 Networking drinks (Asser Institute)

THURSDAY 14 FEBRUARY 2019 – AI AND RESPONSIBILITY

9:30 - 10:30 **The Challenge of Accountable AI**
Dr Machiko Kanetake (Utrecht University)

10:30 - 11:00 Coffee/Tea Break

11:00 - 12:00 **AI and Individual Criminal Responsibility**
Dr Marta Bo (Asser Institute; IHEID) and Abhimanyu George Jain (IHEID)

12:00 - 13:00 Lunch

13:00 - 14:00 **AI and State Responsibility**
Dr Berenice Boutin (Asser Institute)

14:00 - 14:30 Coffee/Tea Break

14:30 - 15:30 **The Morality of Artificial Agents**
Professor Massimo Durante (University of Turin)

15:30 - 16:00 Coffee/Tea Break

16:00 - 17:00 **Group Debate: Towards a Framework of Shared Responsibility for and of AI**
Moderated by Dr Berenice Boutin (Asser Institute)

FRIDAY 15 FEBRUARY 2019 – AI GOVERNANCE

9:30 - 10:30	**The Geopolitics of AI** *Dr Tim Sweijs (The Hague Centre for Strategic Studies)*
10:30 - 11:00	Coffee/Tea Break
11:00 - 12:00	**AI and Political Theory** *Dr Haye Hazenberg (TU Delft)*
12:00 - 13:30	Lunch
13:30 - 14:30	**The Role of International Norms and Institutions in the Governance of AI** *Irakli Beridze (UNICRI)*
15:00 - 17:00	Closing ceremony and high tea reception (Asser Institute)

The 2019 Winter Academy on Artificial Intelligence and International Law was organized with the support of the Municipality of The Hague