HCSS Security

Artificial Intelligence and Its Future Impact on Security

*Testimony prepared by **Dr. Tim Sweijs** for The Committee on Foreign Affairs and the subcommittee on Security and Defense of the European Parliament*

Bruxelles, 10 October 2018

Dear Members of the Committee on Security and Defense,

Thank you for hosting a discussion on a topic that is so tremendously important to the security of Europe and European citizens in the years and decades to come.

And thank you for inviting me to share my thoughts with you.

The past few years there has been rapid progress in various fields that are generally identified under the header of Artificial Intelligence or AI. That progress has largely taken off because of advances in deep learning based on neural network pattern recognition. These advances have been driven by a combination of massive investment from predominantly private actors in AI, persistent increases in computing power, and the availability of large datasets.

AI is predicted to radically disrupt and transform industries, labor markets and societies. Examples that we already see today include the automotive industry (think of driverless cars), health sciences (advanced melanoma screening and detection), financial trading (algorithmic asset management), and advertising (behavioral targeting).[1]

AI is also expected to have a profound impact on the character and perhaps even the nature of future conflict. Leading military powers such as the US, China and Russia, are currently investing heavily in AI related defense R&D efforts to increase their military capabilities. The goal of these investments is to improve the effectiveness and efficiency of current-generation military capabilities, but at the same time they are also intended to bring about disruptive change, to give these powers a competitive edge in military terms. The security impact of AI is an issue of grave concern that is fundamental to future international peace and security.

Today I will start off with a few introductory remarks on what AI is, and what AI is not, to clarify any potential misconceptions that may exist. I will offer a framework to think about AI and its future development, and I will identify five important issues to consider the implications of AI in the context of security and defense.

I will then briefly turn to US, Chinese and Russian efforts in this area, and touch on some AI security and defense applications that defense organizations are currently working on. I will conclude with a brief assessment of the potential positive and negative security implications of AI.

What AI is, and what it is not

We may be currently still on the ascending slope of the AI hype cycle, but the fact that there is so much to do about AI is definitely *not* without reason. Just to give you a few examples related to recent progress:

In the past three years alone, AI systems have proven that they can not only meet but also exceed human performance in image recognition, speech transcription and direct translation. AI systems have learned how to drive; identify relevant information in a paragraph; recognize human faces (even if pictures are blurred) as well as human emotions; AI system have created their own encryption schemes and detect malware; diagnose crop diseases; write cookbook recipes, sports articles, movie screenplays, music and poetry; and AI systems have also beaten world champions at strategic games such as Chess, Poker and Go.[2]

These are all activities that we humans, we Homo Sapiens, or “wise men” and “wise women”, associate with intelligence, and intelligence we associate or at least we used to associate with humans. This is why we are now attaching the label artificial to this newly emerging form of intelligence.

If we want to have a fruitful discussion on what AI means for security and defense, we first have to define it, and describe what we mean by it. So please bear with me for the next couple of minutes.

First: What is intelligence? Intelligence has been usefully defined by Stanford University’s Formal Reasoning Group, as “the computational part of the ability to

achieve goals in the world."[3] This definition of intelligence refers to internal processes ("computation…") that act in the service of bringing about results ("…the ability to achieve goals…") across complex environments ("…in the world"). This definition, as you can see, can be applied both to human and artificial intelligence.

And yes, if we turn to Artificial Intelligence, prevailing definitions typically denote a wide range of capabilities covered by human intelligence that include these elements. AI is about "[the study of the] computations that make it possible to perceive, reason, and act".[4] In a defense context, the US Defense Science Board simply states that AI is "The capability of computer systems to perform tasks that normally require human intelligence (e.g., perception, conversation, decision-making)".[5] Together these provide the relevant building blocks of a definition of AI.

A lot of the progress in recent years comes from new approaches to machine learning – helping an AI system learn to identify deep, hidden patterns in existing datasets. Machine learning, in turn makes uses of a large and growing number of algorithms, which are employed by different approaches to Machine Learning. We do not need to go into these algorithms here, as long as it is clear that AI is not one abstract entity, one SuperComputer, one Killer App, let alone a Sentient Being, but that AI is rather manifested in a range of applications, that each can serve different functions.

There is one additional distinction that is particularly useful when thinking about AI and future progress: Artificial Narrow-, General-, and Superintelligence.

- Artificial Narrow Intelligence (ANI or "narrow AI"): ANI refers to machine intelligence that equals or exceeds human intelligence for specific tasks. Existing examples of such systems are IBM's Deep Blue (chess) and Watson ('Jeopardy!'), Google's AlphaGo (Go) or high-frequency trading algorithms (Wall Street).

- Artificial General Intelligence (AGI or "strong AI"): machine intelligence meeting the full range of human performance across any task;

- Artificial Superintelligence (ASI): machine intelligence that exceeds human intelligence across any task.[6]

We're currently mostly at ANI, with ANI-applications starting to overtake human performance in narrow domains but we're far from a situation in which AGI, let alone ASI is near. Naturally, a critical question concerns future progress in AI: when will AGI and ASI arrive? Expert opinion is divided on this, but a majority of experts expects that AGI is not likely to arrive at the scene in the next two decades. Please also note that there are experts that expect this to be much sooner. Ray Kurzweil, Google's Director of Engineering, for instance, last year moved forward his prediction of when the 'singularity' will take place to 2029.[7]

Even absent a breakthrough from ANI to AGI, there is a lot of 'low hanging fruit' out

in the form of algorithms that have yet to be integrated in existing processes both in the civilian and the military arena. That's why many analysts, myself included, expect progress in AI to continue, also because of the enormous amounts of financial and intellectual investment by large private actors in a whole range of industries complemented by public actors, considerable breakthroughs in hardware, and ever bigger data that become available.

Now: there are five key points I'd like you to take away from this that are relevant to understand and talk about AI in a security and defense context:

First, similar to Human Intelligence, Artificial Intelligence is not one thing, not one entity, not one specific capability. Instead, AI is an enabler. AI can encompass a wide variety of functions in a variety of applications that involve image and sound recognition, knowledge representation, deduction, reasoning and problem solving, and planning, and, when coupled with physical objects, also execution.

Second, these applications can be put to work for good and for bad, especially in the context of defense and security. AI applications can help create better early warning models for the onset of conflict so decision makers can move from early warning to early action and prevention; AI applications can help identify and deradicalize persons with extremist ideas as Google did with its Jigsaw project; AI applications can help save lives of soldiers through better on site medical diagnoses; AI can put fewer of them in harm's way in the first place because of unmanned, which is different from autonomous, systems; and AI applications can automatically patch up cyber vulnerabilities in critical vulnerabilities in minutes where a human team would take months.

However, AI applications can also be used in an offensive way precisely to identify and exploit such cyber vulnerabilities; AI applications can allow actors to provide defensive and offensive systems with greater degrees of automation and autonomy; some conflict actors may eventually grant full autonomy to offensive systems by taking the human out of the loop entirely; and, even absent that situation, AI applications can provide a critical edge to military organizations, who can observe, orient, decide and act much faster, as a result of which existing military capabilities may gradually lose their value. Finally, the integration of AI applications can inject greater uncertainty and friction in the security environment, and diminish predictability, which is likely to aggravate the odds of escalation.

Precisely because AI application can be used for good and for bad, we need to approach the emergence of AI with caution at the same time as we need to carefully distinguish between different uses, purposes and applications in considering its impact.

Third, these functions are executed through algorithms that run on software. Most, although certainly not all, of these algorithms and the software they run on, is of a

dual use nature, which means that it can be used both in military and in civil contexts. Software, as you know, is eminently scale-able: it is essentially an issue of copying the code and you can then apply that software program elsewhere. Therefore, once you have a specific AI application working, it is comparatively easy and comparatively cheap to roll that out more widely in similar systems, certainly compared to the cost of traditional military hardware, where the production of each platform costs a lot of money.

Fourth, precisely because such AI applications are eminently scalable, there are significant implications for the question how the production, the proliferation and the use of a wide variety of AI applications can be controlled by international regimes. In the nuclear domain, for instance, nuclear weapon powers require an entire infrastructure to sustain their capability which includes delivery vehicles, testing and enrichment facilities, critical technologies, and a supply chain, which can be monitored. In the AI domain that is not necessarily the case. This situation is further complicated by the fact that a lot of AI research is done by private actors, who are not necessarily under government control or willing to work with governments. How to control the production, proliferation and AI, is the topic of the second session this morning and this will therefore be useful to discuss in that context.

Fifth, near term development in AI is likely to take place in ANI or Artificial Narrow Intelligence. Further, more disruptive developments, are likely to be farther out in the future. But also in the coming decade, plenty of ANI applications are likely to be very important for defense organizations and will affect the way they operate as well as fight.

So, AI is *multifaceted*, it is *multi-purpose*, it is *scalable*, it is not necessarily easily *controllable*, and AI is likely to have both *long term* but also *short term effects* that are respectively *disruptive* and *incremental* in nature.

AI activities in security and defense today

Now in a security and defense context, the potential impact of AI has not escaped the attention of military organizations. In fact, the defense organizations of leading military powers the US, China and Russia are taking AI very seriously and are expending serious efforts in exploring the potential of both piecemeal and disruptive effects of AI. These three states have announced AI policies and strategies, they have established units at the heart of their defense departments, and are investing literally billions and billions of *euros.*

We don't have enough time to go into the specifics in a lot of detail, but just to give some illustration to these general observations: in 2014 the US launched the Third Offset Strategy with the specific objective to preserve its military pre-eminence through breakthroughs in AI enabled areas including (1) Autonomous Deep Learning

Machine Systems (2) Advanced Human-machine combat teaming and (3) Network-enabled semi-autonomous weapons. Only this summer the Pentagon, established the Joint Artificial Intelligence Center (JAIC) with a budget of US$ 1.7 billion budget over the next five years. This month the US Defense Advanced Research Projects Agency announced the AI Next campaign with a budget of US$ 2 billion to kickstart the next wave of AI innovation. In short, US strategists and policymakers really think that AI will provide the US with a military edge, and they are not only talking the talk, but also walking the walk where it concerns AI related defense R&D.

For China a similar picture emerges. Last year, China announced its plan to be a world leader in AI in 2030. The objective of the People's Liberation Army is to make the shift from 'informatized' to 'intelligentized' warfare. In order to achieve this, China has established the Intelligent Unmanned Systems and Systems of Systems Science and Technology Domain Expert Group. It has also set up a fusion military-civilian AI cell, in order to facilitate the military uptake of civilian AI R&D conducted by private companies. Russia is also committed to developing its AI capabilities. From Vladimir Putin we have the famous quote that "whoever becomes the leader in this sphere will become the ruler of the world." Russia has established the Russian Foundation for Advanced Research (RFAR) in order "to close the gap in advanced research between Russia and the West." It is especially seeking to catch up in the development of autonomous systems. The Russian military industrial committee set the ambitious target of making 30% of military equipment robotic by 2025.[8]

While Russia is often described as very much focusing on the development of autonomous and semi-autonomous systems, for the US and China, it seems that efforts target the entire observe-orient-decide-act or OODA Loop. AI applications that are being researched and/or developed concern early warning models for conflict onset; intelligence collection and analysis; battle space analysis; target acquisition; medical diagnosis; real time translation; automated cyber operations; unmanned semi-autonomous systems; new swarming concepts with large numbers of small, disposable vehicles; and also, and this is very serious, AIs that can explain themselves to humans to improve trust between humans and machines.

Implications

A key question concerns how the emergence of AI will affect global security. This is early days and AI will have myriad impacts. Let me highlight only a few here, both on the positive and the negative side.

On the positive side, more precise targeting through AI, may make war – paradoxically - more humane. Similarly, there may be fewer fatalities as a result of man-machine teaming and uniformed personnel not being put in harm's way. AI enabled better medical treatment is likely to result in lower battle death rates. Similarly, AI will open up opportunities to improve early warning to inform early

action, because it will provide better and more timely insight in causes of conflict and escalation processes, and may even facilitate de-escalation in strategic decision making processes because AIs will be able to furnish different options than commonly considered.

On the negative side, there are plenty of serious negative implications too. The inclusion of AI in existing military capability portfolios can upset the existing military balance of power by making traditional systems and doctrines obsolete. The political costs of going to war are likely to be lower if fewer humans are sent to front thereby reducing an important constraint to go to war. Immature AI applications rushed into military platforms can lead to spiral dynamics when they run out of control in uncertain environments, as sometimes happens in financial trading. What is more, conflict can speed up to such a degree that human intelligence can no longer keep up especially if some conflict actors decide to remove people out of the loop and leave the decisions to AI applications. This will provide their opponents with an important dilemma whether to keep the human in the loop and possibly face defeat or follow suit. When that happens, a battle between AIs may unfold, and then not only the character, but also the nature of war will change, leading to a whole new era in human history.

That, Ladies and Gentlemen of the Commission, remains science fiction for now, but that scenario does vividly illustrate that very important ethical and legal questions can never be considered in isolation from the military-strategic ramifications associated with the rise of new military technologies. I am therefore pleased that we are talking about this today and I look forward to your questions and comments.

Thank you.

Dr. Tim Sweijs is Director of Research at The Hague Centre for Strategic Studies (HCSS).

Endnotes

[1] For the examples, see Stephan de Spiegeleire, Matthijs Maas, and Tim Sweijs, *Artificial Intelligence and the Future of Defense: Strategic Implications For Small- and Medium-Sized Force Providers* (The Hague Centre For Strategic Studies, 2017), 22, https://hcss.nl/sites/default/files/files/reports/Artificial%20Intelligence%20and%20the%20Future%20of%20Defense.pdf.

[2] As we describe in de Spiegeleire, Maas, and Sweijs, 44.

[3] John McCarthy and Stanford University Formal Reasoning Group, "What Is Artificial Intelligence | Basic Questions," Formal Reasoning Group, 2007, http://www-formal.stanford.edu/jmc/whatisai/node1.html

[4] Patrick Henry Winston, Artificial Intelligence, 3rd ed., 1992.

[5] Defense Science Board, "Report of the Defense Science Board Summer Study on Autonomy" (Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, June 2016), https://www.hsdl.org/?view&did=794641, p. 5

[6] See de Spiegeleire, Maas, and Sweijs, *Artificial Intelligence and the Future of Defense: Strategic Implications For Small- and Medium-Sized Force Providers*, 30.

[7] See de Spiegeleire, Maas, and Sweijs, 50–51.

[8] For more background, see de Spiegeleire, Maas, and Sweijs, 77–87.