Forum: The North Atlantic Treaty Organization (NATO)
Issue: Protecting Against Emerging Threats Posed by Artificial Intelligence (AI)
Names: Bene Voranan Puengchanchaikul and Chaarvi Mehta
Positions: Head Chair and Deputy Chair

Introduction:

The increasing prevalence of Artificial Intelligence (AI) is undeniable in today's societies. It has completely transformed–and will continue to transform–the way NATO operates in regards to international security. As machines begin to be able to perform tasks that normally would call for human intelligence, AI presents some notable threats and challenges in the way NATO addresses its core tasks of collective defence, crisis management, and cooperative security in terms of both traditional military capabilities and multinational threats.

The Allied Defence Ministers met in October 2021 and formally adopted the NATO Artificial Intelligence Strategy. According to this strategy, NATO and its allies must make certain that any and all AI applications developed and reviewed for deployment are in line with the organisation's six Principles of Responsible Use (PRUs). These principles include lawfulness, responsibility and accountability, explainability and traceability, reliability, governability, and bias mitigation.

On Tuesday, February 7th, 2023, NATO's Data and Artificial Intelligence Review Board (DARB) held a meeting to discuss and establish "a user-friendly and responsible Artificial Intelligence (AI) certification standard". In theory, the aim was to help allies' industries and institutions keep AI and data projects compliant with both international law and the common norms and values of NATO.

However, since these AI guidelines are still relatively new, concerns and threats may arise. To begin with, NATO has to realistically coordinate the overseeing of AI

development and application in defence on a national level. This means dealing with different countries' regulations in terms of usage, accessibility, sharing, and transfer of technology. While some allies believe in emphasising the sharing of data in addition to practical guidelines for AI-enabled systems in terms of various operational uses, others are more sceptical about the lack of detail in the six PRUs. Moreover, some members believe that they will be giving up technological benefits to competing nations by having such a standardised AI-regulation system.

In conclusion, AI is so new that laws regulating AI development and implementation are still in the works. Even in countries that are leading AI developers, a clear, AI-specific act is still not present to address growing threats that the innovation poses.

Definition of Key Terms:

1. **Artificial Intelligence (AI):** (noun) the theory and development of computer or technological systems that can complete tasks that normally call for human intelligence by enabling these systems to think like humans, which includes skills such as but not limited to visual perception, speech recognition, decision-making, and language translation
    a. **Artificial Narrow Intelligence (ANI) or "Weak AI":** (noun) AI that can complete specific tasks without the involvement of general cognitive abilities
    b. **Artificial General Intelligence (AGI) or "Strong AI":** (noun) AI that imitates human intelligence and can understand, learn, and apply knowledge across different fields
    c. **Artificial Superintelligence (ASI) or Hypothetical AI:** (noun) Ai that can potentially solve complex problems and make improvements beyond human intelligence and comprehension
2. **Cooperative security:** (noun) activity among two or more states with the goal of reducing the risk of war and the consequences of a war as well as ensuring the

overall safety of the states without directly targeting a specific state or group of states

3. **Cybersecurity:** (noun) the state of and methods towards protection against the criminal or unauthorised usage of electronic data, including policies, training, steps, etc. that protect the virtual environment
4. **Governability:** (noun) the extent to which a social being or system can be governed (managed; controlled)
5. **Intellectual property (IP):** creations of the mind, including but not limited to literary and artistic works, designs, symbols, names, and images
6. **Oversight:** (noun) the action of overseeing or supervising something

Background Information

Artificial intelligence

Artificial intelligence was first found in 1956 Dartmouth College. AI–defined as “a set of sciences, theories and techniques” which has the goal of imitating a human’s cognitive abilities–has existed as a branch of knowledge for over 60 years. AI can complete skills such as but not limited to mathematical logic, statistics, probabilities, computational neurobiology, and computer science. Because of the development in computers and the access to large amounts of data, the world has seen a dramatic rise in exposure to and use of AI starting in 2010.

1940-1960

From 1940 to 1960, technological developments occurred at a great pace, including being accelerated by World War II. According to American computer scientist Marvin Minsky from Carnegie-Mellon University, AI is "the construction of computer programs that engage in tasks that are currently more satisfactorily performed by human beings because they require high-level mental processes such as: perceptual learning, memory organisation and critical reasoning.” In contrast from today, the early 1960s saw a fallback of the popularity of technology, which included less advanced machines than at present.

1980-1990

At this time, the launch of the development of AI expert systems was sent into motion. This started with MIT's DENDRAL, an expert system specialised in molecular chemistry, in 1965 and Stanford University's 1972 DENDRAL, which specialised in the diagnosis of blood diseases and the prescription of drugs. Both systems integrated an "inference engine". This engine was made to logically mirror the reasoning of a human. So, these machines were able to output answers with a high expertise level when data was inputted. Nevertheless, the focus on AI and these types of machines retreated once again towards the end of the 1980s and the start of the 1990s. The issue was that the development and maintenance of these technologies had become very difficult. Plus, there were still other methods to complete tasks that were faster, less complex, and less expensive than the expert systems. As a result, in the 1990s, the term "artificial intelligence" had neared being a taboo as the sudden shifts and innovations worried many people.

Since 2010

Since around 2010, the AI field–research, development, and implication–has seen rapid growth. Two main factors are responsible for this.

First, this past couple of years has experienced a huge increase in accessible data in more and more subject areas. An expanded data bank means a wider range of information readily, easily available for both algorithms and for the public–like a simple Google search that grants access to millions of data online, for example.

Second, a key factor in the emergence of modern AI is the discovery of computer graphics card processors. These cards are increasingly efficient and have helped "accelerate the calculation of learning algorithms" in AI. Before 2010, processing a sample or data–a lengthy and repetitive process–took a very long time. However, at present, these cards' computing power at over a thousand billion transactions per second allows us to make much progress in terms of AI development at less of a financial cost.

Although these newly introduced technologies have introduced things such as text recognition, the development and use of AI still proves to be controversial and challenging to navigate. People still pose the following question: Will AI truly be able to replicate human interactions, thinking, and analysis? If so, will it come at a higher cost than the benefits it provides?

Current Situation

Artificial Intelligence (AI) is and will continue to push boundaries and create new obstacles when it comes to information security. Although some may be flawed and pose threats, it is accurate to say that these systems will continue to independently learn, reason and act, replicating human behaviour closer and closer.

Currently, AI presents information risks like never before. AI also makes the flow of information and data that had existed beforehand even more dangerous. On the other hand, AI can play key roles in organisations' defensive arsenals as well. Hence, prior to these technologies being embraced as a vital part of everyday business, both businesses and information security leaders have to understand its risks and opportunities

At present AI has weaved its way into many daily-life tasks. For instance, multiple variations of AI are used by organisations to assist in areas such as customer service, human resources and bank fraud detection. Even so, the prevalence of AI can result in confusion and scepticism over what AI actually is and what its role means for businesses and security.

Threats due to AI are even more prominent and concerning when it comes to AI possibly replacing expert security and medical practitioners. So, it's necessary to establish a balance between the need for human supervision with the confidence to let AI take control independently and effectively. This will take time to perfect, considering mistakes and poor decisions made by these still-developing machines, so it's crucial for states to provide a framework to guide AI towards the right direction.

When organisations adopt defensive AI, experts must put in time, training and support in bridging the work done by humans and the work done by AI systems. Once

this is achieved, human and artificial intelligence combined has the capability to be a valuable component in the defences of an organisation, community, nation, or region.

Major Parties Involved and Their Views

Canada

Canada, a member of NATO, is a leader in AI development, focusing on AI implications in healthcare, education, airlines, and entertainment. Canada faces fundamental risks with AI such as its lack of AI regulation and oversight. Nevertheless, Canada continues to look at AI that can improve daily life, such as AI in school classroom settings. However, these developments must be made with caution as there are still many persistent threats presented in Canada's AI. The Canadian federal Integrated Terrorism Assessment Centre's May 2023 report via the Access to Information Act says that "AI-generated hoaxes pose a 'persistent threat' to public safety".

United States of America

As a member of NATO, the United States's priority is making sure that allies focus on agreeing to "practical guidelines for the operational use of AI-enabled systems". It believes that data should be shared in order for this to happen. The United States is a big supporter of AI, having contributed an increasing amount of funds for its artificial intelligence companies from 300 million in 2011 to 16.5 billion USD in 2019. The U.S. government aims to utilise AI in healthcare, transportation, the environment, and benefits delivery. Additionally, in 2022, the US passed its Algorithmic Accountability Act. This act calls for all "companies to assess the impacts of the automated systems they use and sell, creates new transparency about when and how automated systems are used, and empowers consumers to make informed choices about the automation of critical decisions" wyden.senate.gov.

Denmark

Another NATO ally, Denmark has published “The National Strategy for Artificial Intelligence”, which is a framework that will potentially allow the country to “be a front-runner in responsible development and use of AI”. The goal of this is to use AI to benefit individuals, businesses, and society. This strategy will allow businesses, researchers, and public authorities to utilise AI to its greatest potential while still being responsible in its development and implication–an example of a strategy that can be expanded past just the national level. More specifically, the strategy outlines four key objectives for the Danish to follow when developing and using AI. First, “Denmark should have a common ethical and human-centred basis for AI”. Second, the people researching and developing AI should be Danish researchers. Third, “Danish businesses should achieve growth through developing and using AI”. Fourth, quality public services should be offered by having the public sector use AI.

Germany

In 2023, Germany had only 42 AI start-ups fail–which is considered a very high survival rate in comparison to other regular start-ups. Recently, in order to boost investments and keep up with China and the US which have both been rapidly developing AI, Germany has announced the launch of its AI action plan. This action plan consisted of an outline for 12 areas for action by the ministry. For instance, “strengthening the entire AI value chain at the national and EU levels, focusing on connecting the dots with education, science and research”. In order to achieve this, €1.6 billion has been pledged by the ministry towards AI investments. These funds will go into promoting 20 AI initiatives as well as 50 ongoing measures for research, skills, and infrastructure development.

UN Involvement, Relevant Resolutions, Treaties and Events

NATO Defence Ministers endorsed a strategy in February 2021. This strategy addressed emerging and disruptive technologies (EDTs) in order to guide allies’ AI development. Since then, NATO has backed multiple strategies to publish two annual reports and participated in multiple discussions on this subject area, including in the

NATO Summit in Brussels in 2021. There is also a NATO Advisory Group on Emerging and Disruptive Technologies. NATO also works in close collaboration with the UN and the European Union (EU).

- Recommendation on the Ethics of Artificial Intelligence, November 2021 **(SHS/BIO/REC-AIETHICS/2021)**
    - Passed by the United Nations Educational, Scientific and Cultural Organization (UNESCO), this resolution was “the first-ever global standard on AI ethics”. All 193 Member States adopted this framework. An important quality that this recommendation focuses on is the protection of human rights and dignity in AI development and implications, combatting any threats to human rights and dignity. It emphasises transparency, justice, and human oversight. A strong point of this recommendation was its “extensive Policy Action Areas”. This component lets policymakers implement the recommendation’s core values and principles while still adhering to the conditions, data governance, environment, ecosystems, gender, education, research, health, and social well-being of each specific country.
- AI for Good Global Summit, Geneva, 6–7 July 2023
    - Also known as the AI Summit, it has been hosted by the International Telecommunication Union (ITU) since 2017. This international AI event has resulted in the launch of over 150 projects. It involves major stakeholders such as governments, industries, academia, media, 40 UN agencies, the American Society for Compute, etc. The event is an action-based platform that focuses on positive AI implications in “health, climate, gender, inclusive prosperity, sustainable infrastructure, and other global development priorities”, according to its website.
- Impact of rapid technological change on the achievement of the Sustainable Development Goals and targets, 26 November 2018 **(A/RES/73/17)**
    - This resolution called upon key groups including the Technology Facilitation Mechanism, the Commission on Science and Technology for Development, and the Economic and Social Council (ECOSOC), asking

these groups to *consider* the rapidly changing technologies–including AI–in the path towards achieving the SDGs. However, the impact was minimal due to a number of the clauses simply asking groups to consider previously made points or do research rather than taking actionable steps. For instance, asking organisations "to continue to consider, in a coordinated manner within their respective mandates and existing resources, the impact of key rapid technological changes, such as artificial intelligence, among others, on the achievement of the Sustainable Development Goals and targets … ."

Possible Solutions

Solutions proposed by delegates should answer the following questions and/or be generated based on the following:

- How can this be practically implemented in countries with different levels of AI development and usage?
- How will this affect AI developers and industries?
- Does this align and expand upon NATO's six PRUs?

Although they must be elaborated on in resolutions and debate, possible solutions include but are not limited to the following:

- Establishing a framework or guideline for states' development of AI
- Requiring all AI systems to go through an analysis or test that will ensure that the AI system will be used responsibly and that it was created with responsible intentions
- Fostering communication between AI developers and the state or government
- Establishing or improving upon clear copyright laws specifically geared towards IP of content generated using AI
    - Requiring users to give credit to AI for any content generated by AI
    - Allowing the copyright to be given to the AI developer but the ownership to be given to the user for their imputed data

    - Establishing a central AI detection system
    - Making clear the rights of the AI developer
- Requiring any and all AI systems to make transparent their sources of data and how data used was collected
- Establishing and communicating clear boundaries
    - What makes AI usage either ethical or harmful/malicious
    - What would be considered malpractice or misconduct

Bibliography

Useful Links

- https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html
    - This source provides an overview of the role of AI in NATO as well as the "Artificial Intelligence Strategy for NATO" adopted by the Allied Defence Ministers at their October 2021 meeting.
- https://www.iiss.org/ja-JP/online-analysis/military-balance/2021/11/algorithmic-power-nato-and-artificial-intelligence/#:~:text=NATO%20defence%20ministers%20have%20formally,%2C%20governability%2C%20and%20bias%20mitigation.
    - This article provides a critique of "the Alliance's first artificial intelligence (AI) strategy", which is helpful in pointing out the possible threats due to the introduction of AI which must be considered by delegates.
- https://www.youtube.com/watch?v=ecEW-ob896w
    - The video linked above highlights the more extreme concerns over AI threatening society, including a possible AI arms race, as well as the importance of collaboration in regulating AI development and implementation.

Works Cited

- https://languages.oup.com/google-dictionary-en/
- https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html
- https://www.nato.int/cps/en/natohq/news_211498.htm#:~:text=At%20present%2C%20NATO%20is%20piloting,climate%20change%20and%20imagery%20analysis.
- https://www.iiss.org/ja-JP/online-analysis/military-balance/2021/11/algorithmic-power-nato-and-artificial-intelligence/#:~:text=NATO%20defence%20ministers%20have%20formally,%2C%20governability%2C%20and%20bias%20mitigation.
- https://www.nato.int/cps/en/natohq/official_texts_208374.htm
- https://www.tandfonline.com/doi/abs/10.1080/13876980802028107#:~:text=Governability%20can%20therefore%20be%20defined,societal%20entity%20or%20system".
- https://www.marshallcenter.org/en/publications/marshall-center-papers/cooperative-security-new-horizons-international-order/cooperative-security-theory-practice#:~:text=Cooperative%20security%20is%20thus%20one,formation%20of%20the%20NATO%20alliance.
- https://www.wipo.int/about-ip/en/#:~:text=Intellectual%20property%20(IP)%20refers%20to,and%20images%20used%20in%20commerce.
- https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/what-is-artificial-intelligence#:~:text=Artificial%20Narrow%20Intelligence%20(ANI)%3A,domains%2C%20similar%20to%20human%20intelligence.
- https://www.statista.com/statistics/672712/ai-funding-united-states/
- https://ai.gov/ai-use-cases/#:~:text=The%20 federal%20 government%20 is%20 leveraging,doesn't%20violate%20their%20rights.
- https://www.coe.int/en/web/artificial-intelligence/history-of-ai#:~:text=The%20summer%201956%20conference%20 at,the%20founder%20of%20the%20discipline.
- https://www.freecodecamp.org/news/how-to-use-artificial-intelligence-in-cybersecurity/#:~:text=Incorporating%20human%20oversight%20in%20critic

al,based%20on%20machine%2Ddriven%20decisions.&text=practices%2C%20and%20threat%20intelligence%20can,against%20malicious%20AI%2Dbased%20attacks.

- https://www.securitymagazine.com/articles/91242-reducing-the-risks-posed-by-artificial-intelligence
- https://www.cna.org/our-media/newsletters/ai-and-autonomy-in-russia
- https://www.nato.int/cps/en/natohq/topics_184303.htm
- https://www.ctvnews.ca/sci-tech/ai-generated-hoaxes-pose-a-persistent-threat-to-public-safety-intel-analysis-1.6685125
- https://en.digst.dk/strategy/the-danish-national-strategy-for-artificial-intelligence/
- https://www.euractiv.com/section/artificial-intelligence/news/german-ai-action-plan-the-answer-to-chinese-and-us-dominance/
- https://www.wyden.senate.gov/imo/media/doc/2022-02-03%20Algorithmic%20Accountability%20Act%20of%202022%20One-pager.pdf
- https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
- https://documents-dds-ny.un.org/doc/UNDOC/GEN/N18/403/98/PDF/N1840398.pdf?OpenElement
- https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202001gls.html
- https://aiforgood.itu.int/summit23/