What Does Resilience-Building to Emerging and Disruptive Technologies Actually Look Like? A Study Addressing the Public Policy Challenges and Socio-Political Implications of the Development of Artificial Intelligence for NATO Security and Defense in Continental Europe

Kulani Abendroth-Dias, PhD student, Graduate Institute of International and Development Studies

kulanidias@gmail.com; Twitter: @kulaniadias

Table of Contents

Understanding the problems

01 The need to build resilience to AI and ML-driven technologies is often discussed in policy circles. However, little attention has been paid to operationalising resilience-building to the rapid development of AI- and ML-driven technologies across the continent.

02 At its core, resilience requires inclusive and forward-thinking regulation and research that can build flexibility to respond to evolving security challenges. Developing resilience to AI- and ML-driven attacks is a central tenet to building citizen trust.

03 How do current policymakers, industry professionals, academics, and non-profit agents operationalize resilience?

Overview of the study

Exploratory, inductive approach

Semi-structured interviews conducted virtually due to Covid-19

Content analysis (Braun & Clark, 2006) conducted to identify themes; frequency analysis

Collate views of policymakers, industry professionals, non-profit agents, and academics within one study to understand overlap and discrepancies

Sample

- Sample: 22 participants virtually interviewed between February - May 2020
- Based in Austria, Belgium, Germany, the Netherlands, Poland, Sweden, Switzerland, and the United Kingdom
- Participants from the United Nations Institute for Disarmament Research (UNIDIR), the International Panel on the Regulation of Autonomous Weapons (iPRAW), the German Council on Foreign Relations, the German Federal Foreign Office, the German Army (Bundeswehr), the Belgian Royal Military Academy, the University of Namur, the University of Siegen, the Free University of Brussels (VUB), Central European University, ETH Zurich, the Hague Center for Security Studies, Djapo, McKinsey and Company, Compagnie Européenne d'Intelligence Stratégique Sprl (CEIS), the Center for Data Innovation, The Democratic Society, Statewatch, Transparency International, the Stockholm International Peace Research Institute, the European Commission, and the European Parliament. academic, policy-making, non-profit, and industrial sectors respectively

Interview Questions (part 1)

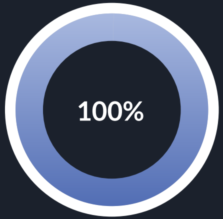
- Please tell me a little bit about your current work.
- Where, to your knowledge, are the physical hubs for developing AI for European security and defense?
- What, to your knowledge, are the different types of AI-driven technologies that are being developed for European security at the moment? Can you speak of any policy or industrial priorities?
- There is evidence to support that risk-averse countries, especially those attempting to avoid military casualties, will continue to support the development of unmanned aerial vehicles to be deployed into battle. Detractors of this argument posit that this will simply escalate arms races and military budgets, and that wars will only be concluded with the loss of human life, whether or not they have to invade civilian territory outside of battlegrounds to do so. What are your thoughts on this debate?
- What, in your opinion, are some, if any, of the challenges to policy-makers given the rapid development of AI for European/NATO security and defense?
- What are some ways you know that policymakers and industrial organizations are collaborating to mitigate the risks of AI-driven forgery, and innovate for European security and defense?
- To your knowledge, are there any counter-AI strategy development organizations, either in policy or industry? Any organizations working to build resilience to machine-learning-driven malignant attacks - either in the cyber world or real-world military (with UAVs, disinformation, etc.)
- In your opinion, what does building resilience to AI-driven attacks look like?

Interview Questions (part 2)

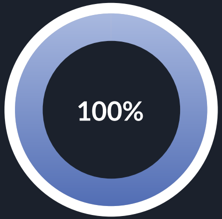
- My next question is on the procurement of AI- and machine-learning driven surveillance systems security and defense. According to a study by the Carnegie Endowment, there are many Chinese and US-based companies pitching these systems to bolster European border defenses, predictive policing, safe cities, facial recognition etc. A breakdown of military expenditures in 2018 shows that forty of the top fifty military spending countries also have AI surveillance technology, including Germany, France etc. What can you say about the current state of AI-driven surveillance technologies being used in Europe, in terms of where it is developed, who it is developed for, and where it is put to use?
- With regard to the debate on the coding of bias into autonomous weapons systems: Human biases can be coded into systems driven by artificial intelligence via biased datasets - what do you think would be the strategic implications of this, and to your knowledge, are there steps being taken to address these biases? If so, to your knowledge, who and how are these steps being taken?
- As you know, defense during the fourth industrial revolution is not just about securing our physical borders. AI and machine learning technologies have been used to target certain segments of the population, mine data, and influence voting behaviors. What more can we do to build resilience to these malignant attacks in your opinion and based on your expertise?
- Is there anything else you would like to add? Any questions you wish I would have asked?

Overall results

Disinformation, privacy protection

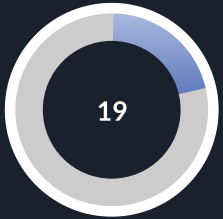


Future wars will not be concluded via the destruction of automated weapons on the field (however developed), but by the death of people, whether civilian or military. Underlined the assistive capacities of AI on the battleground, especially with regard to reconnaissance and carrying out surgical strikes



AI- and ML-driven technologies will contribute to the increase of military budgets and the escalation of arms races

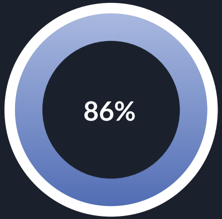
Sharing one's data is the default of the future, and an innovative solution would be a centralised space for every citizen to share what types of data they wish to share with, and restrict from, the private sector and public sector respectively



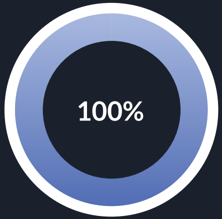
Use of AI in polarising societies via disinformation

Overall results

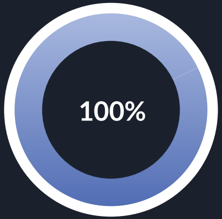
Location of AI hubs



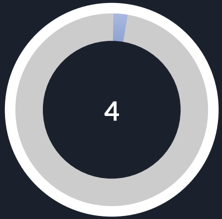
Difficult to identify a "hub" for AI development in Europe per se, stating that AI development is quite diffused across member states



Dual-use nature of AI-driven technologies, with their use in offensive and defensive campaigns, and in civilian and military domains, increased risk in implementing AI-driven solutions developed outside the EU within member states



China and US hubs for AI development



Only four participants mentioned the role of the European Commission, NATO, and the European Defence Fund by name

Germany, France, the UK, Sweden, and the Netherlands were mentioned as areas of AI development, with references to London (91%), Berlin (68% of the sample), Amsterdam (50%), and Paris (27%) made. Also Israel, India, Russia, South Korea

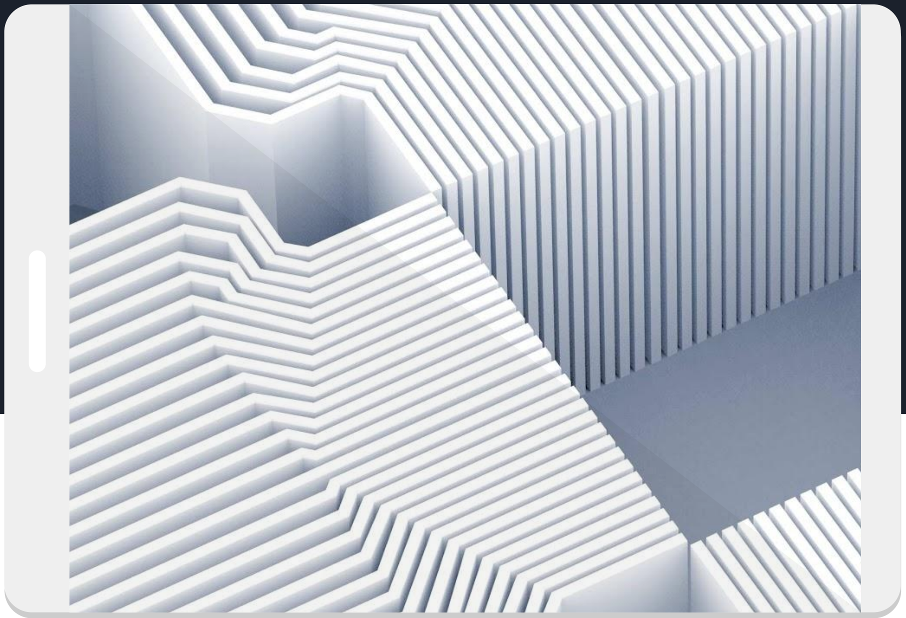
However, which areas are being targeted? How are localised solutions for resilience being developed? NATO carries out defense at its borders

Responsible AI

All participants emphasised the **transnational or international** nature of the European defence industry, making references to collaborative projects in the development of AI-driven solutions.

50% were able to recall the **name of an organisation** working on responsible AI.

References to Statewatch, Transparency International, Algorithm Watch, and the United Nations were made.



In response to the role of **AI in increased surveillance**, 13 participants called on the United Nations or NGOs such as those mentioned above to work with policy-makers to regulate their use.

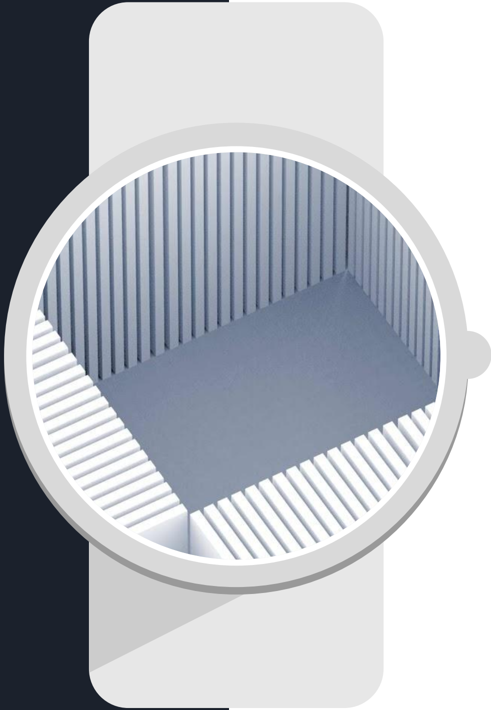
The importance of data protection to combat surveillance was mentioned by **all** participants.

Algorithm and automation biases

All participants noted the importance of **collaboration** across EU member states in developing and regulating AI-driven solutions.

All participants were aware of the risk of coding human biases into ML-driven algorithms, and 50% were able to recall examples of such occurrences.

Five participants recalled the now famous case of facial recognition technology implemented at the Berlin-Südkreuz train station resulting in a 20% error rate, mostly misidentifying people of colour at the station.



Seventeen participants (77% of the sample) mentioned that white males were explicitly not at risk from algorithmic bias.

Two participants identified women as those at risk of this type of bias, whereas 50% of the sample (11 interviewees) noted that people of colour were most at risk of this type of error.

Insights for Policy and Security Professionals

Recommendations mentioned by at least 50% of participants

- Adapting the Operationalisation and Regulation of AI- and ML-Driven Technologies to Existing International Human Law (IHL) Frameworks
- What Does Retaining Human Responsibility for AI-Driven Attacks Look Like?
- Moving Beyond "Meaningful Human Control"
- Legally Binding Instruments to Regulate the Use of AI in International and EU Versus National Contexts
- Critical Thinking Skills Versus Digital Literacy
- Developing AI Competencies and Regulations Within the European Startup Ecosystem
- Increasing AI-Driven Solutions for Military Use
- Understanding the Risks of Not Using AI
- Increasing the Technical Awareness of the Development and Capacities of AI- and ML-Driven Within Policy Solutions
- Developing the Responsible Democratisation of Data

Thank you for your time and attention

kulanidias@gmail.com

@kulaniadias - Twitter

Please note that the contents of this presentation have been carried out in the author and presenter's personal research capacity and do not reflect the official views of any of the organizations with which she is currently affiliated.