THE End of Security as we Know it?

Understanding the Security Challenges of Artificial Intelligence

Workshop and Public Debate
WEDNESDAY, 25 SEPTEMBER 2019

Rhein-Main Universitäten
Eine strategische Allianz

#ChanceKI
wissenschaftsjahr.de

An Initiative of the Federal Ministry of Education and Research

Science Year 2019
ARTIFICIAL INTELLIGENCE

Workshop

Understanding the Security Challenges of AI

Artificial intelligence (AI) is the theme of the Science Year 2019 of the German Federal Ministry of Education and Research. In the context of the increasing integration of AI into the operations of governments and businesses, policy professionals have to cope with potential security challenges of AI.

The event provides a concise overview on the security challenges in the fields of data protection, cybersecurity, law, business, society and international politics.

Senior policy professionals will discuss salient security challenges of AI with experts in three small group discussions of 60 minutes each. After one hour of discussion the policy professionals move to a different table. In doing so they will be acquainted with different perspectives and approaches to the security challenges of AI.

CVs and Take Away Messages of the Participating Experts

Raluca Csernatoni

Raluca Csernatoni is a guest professor at the Institute for European Studies of Vrije Universiteit Brussel and a visiting researcher at Carnegie Europe in Brussels, where she works on European security and defence with a specific focus on disruptive technologies. Previously, she was a postdoctoral researcher and lecturer at Charles University's Institute of Political Studies in Prague. She has also conducted research at the Royal Higher Institute for Defence's Centre for Security and Defence Studies, at the International Security Information Service Europe in Brussels, and as research fellow in the Study Program on European Security at the Institute for European Politics in Berlin. Csernatoni holds a PhD and master's degree in International Relations from Central European University.

Take away messages

(1) We live in times of high-tech euphoria marked by instances of geopolitical doom-and-gloom. There seems to be no middle ground between the hype surrounding disruptive technologies such as Artificial Intelligence (AI) and their impact on security and defence, and anxieties over their potential destructive consequences and misuses.

(2) This emphasis on techno-solutionism in military affairs is nothing new. Predictably, AI is often discussed as a potentially disruptive weapon and likened to prior transformative technologies such as nuclear and cyber, placed in the context of national security. However, this framing is highly problematic and sets the AI's parameters as being one-dimensional.

(3) Equally, the emergence of new technologies such as AI has ushered in widespread and differing assumptions about the future of war and security. Almost every advance in AI is framed and hyped as a revolution, yet not every new technology will fundamentally alter security practices.

(4) In this regard, there is a clear need for 'transdisciplinary' conceptual lenses at the intersection of various academic disciplines to tackle disruption and game-changing security technologies. Such work should focus on examining the assumption of a paradigm shift is warfare and the ways ahead for International Relations and other disciplines to reconceptualise central categories of thinking about security, war, agency, actorness, and human-machine interactions.

Christopher Daase

Christopher Daase is Professor of International Organizations at Goethe University Frankfurt, Deputy Director at the Peace Research Institute Frankfurt (PRIF) and Principal Investigator at the Cluster of Excellence "The Formation of Normative Orders". His main research interests lie in the areas of international security and transnational politics. He is one of the editors of the Peace Report, which analyses current violent conflict, shows trends in international foreign, security and development policy and gives concrete recommendations to the German federal Government concerning peace policy-making.

Take away messages

(1) Artificial intelligence requires a rethinking of security, i.e. international, national, social and human security. AI offers unprecedented opportunities as well as risks for the coexistence of people and states.

(2) AI has the long-term potential to change the power balance between states. There is already fierce competition for the most advanced AI technologies. This competition impedes an orderly, controlled and responsible use of AI technologies in terms of *international security*.

(3) Regarding *national security* AI is applied in the development of autonomous weapon systems, which allow states, but also non-state armed groups, to coordinate larger covert attacks. International cooperation and preventive arms control are urgently required to limit the development of fully autonomous systems.

(4) Social security, i.e. the safeguarding of social and political freedoms vis-à-vis the state, can be drastically impaired by AI. A social dialogue is necessary to assess the chances of crime prevention regarding the risks of total social control.

(5) Humanitarian security, i.e. the global provision of basic human needs, could be strengthened by AI if, for example, crises were predicted in good time and crisis management were optimised. However, this would require AI-expertise in international and humanitarian institutions, which do not render short-term economic or political profit.

Rolf van Dick

Rolf von Dick is professor of Social Psychology at Goethe University Frankfurt and one of the directors of the Center for Leadership and Behavior in Organizations (CLBO), a platform in which colleagues from Economics, Sociology and Psychology study together and in exchange with practitioners. He served as dean of the department of psychology and sports sciences. As vice president of Goethe University, he is responsible for international affairs, PhD and post-docs, and diversity and equality.

Rolf van Dick published more than 200 books and scientific articles and he served as (associate) editor of the European Journal of Work and Organizational Psychology, the British Journal of Management, the Journal of Personnel Psychology, and the Leadership Quarterly.

Take away messages

(1) Technical challenges are going to be solved, some challenges are even going to be solved much quicker than expected. We have to focus on societal challenges, i.e. we have to ask questions in terms of ethical, legal and psychological issues and find the respective answers.

(2) We used to chase technological progress by adjusting man to technology. We will have to think the other way by putting human needs first and adjusting technologies accordingly.

(3) AI will not supplant humans. Algorithms will conduct many tasks including medical diagnosis, legal services or teaching in universities. However, in the future algorithms will be conducting some types of work including highly qualified tasks in medical diagnosis, legal services or teaching in higher education. In return new types of work as well as completely new professions will emerge.

Gianclaudio Malgieri

Gianclaudio Malgieri is a doctoral researcher at the Law, Science, Technology and Society" Research Group of the Vrije Universiteit Brussel, where he is Work Package Leader of the EU H2020 PANELFIT Project. He is also Training Coordinator of the Brussels Privacy Hub and lecturer of Data Protection Law and Intellectual Property for undergraduate and professional courses at the University of Pisa, Sant'Anna School of Advanced Studies and VUB. Gianclaudio is also qualified to practice Law in Italy.

He got an LLM with honours at the University of Pisa and a JD with honours at Sant'Anna School of Advanced Studies of Pisa (Italy). He was visiting student at the OII of the Oxford University, London School of Economics, World Trade Institute of the University of Bern and École Normale Superieure de Paris. He published more than 30 articles in leading international law reviews, and he's also editor of the Italian Handbook of Personal Data Protection.

Take away messages:

(1) Researchers and policy-makers are timidly starting to explore a new dimension of cybersecurity in AI: cognitive security. AI is not just capable of discrimination, but also of manipulation affecting mental integrity of consumers/users/citizens/voters, through hyper-nudge and hyper-personalized behavioural ads.

(2) Usually we consider just "traditional" security risks (availability, authenticity, integrity and confidentiality). With the rise of AI, we should extend the notion of security "risks", and the GDPR helps to do so in two parallel ways:
   a. focussing on risks to "fundamental rights and freedoms" (Article 35 GDPR);
   b. considering the secondary effects of traditional security risks: psychological distress, reputational damage, etc (ENISA guidelines), or any other significant economic or social disadvantage (recital 75, GDPR).

(3) The model of co-governance of security risks prevention in AI and automated data processing should be developed more. The EU data protection framework, for example, creates a binary system based on individual rights and accountability duties. These two levels should be merged in a collaborative governance tool. What can really help is the Data Protection Impact Assessment, especially when applied to Automated-Decision Making system (the so called Algorithmic Impact Assessment).

Kai Rannenberg

Kai Rannenberg holds the Chair of Mobile Business & Multilateral Security (www.m-chair.de) at Goethe University Frankfurt. 2004-2013 he was academic expert in the Management Board of the EU Network and Information Security Agency, ENISA; and since 2013 he is member of ENISA's Permanent Stakeholder Group.

He has been coordinating several leading EU research projects, e.g. the Network of Excellence "Future of Identity in the Information Society (FIDIS)" and the Integrated Project "Attribute based Credentials for Trust" (ABC4Trust). Currently he is coordinating CyberSec4Europe, a pilot for the European Cybersecurity Competence Network the EU is aiming for.

His research interests include mobile and embedded systems and multilateral security in e.g. m-business, LBS, transport systems, and industrial applications; privacy and identity management, especially attribute based authorisation; communication infrastructures and devices, e.g. personal security assistants and services; security and privacy standardisation, evaluation, and certification.

Take away messages and questions

(1) Many security decisions, e.g. on intrusion, attacks, or data leaks, need to be made under time pressure. Decision support systems based on AI concepts are therefore gaining popularity. However such systems still need to improve in the way, they give assurance to their users. At the same time a systematic way to assess systems based on AI is missing. One reason is, that many systems are constantly "learning", i.e. changing their behaviour, which outdates any assessment quickly.

(2) One way to deal with the lack of assurance in AI systems is to limit the learning phase to a trial and learning period and stop the learning and change of behaviour, once the system is installed in a critical environment. Is this viable given today's attacks dynamics? Or is it the only way to avoid being misled and damaged by subverted AI systems?

(3) Some questions to help the assessment of AI systems their algorithms may be:
   a. How to make what looks like black-box-decision-making trustworthy to stakeholders?
   b. How to best assess systems for biased decisions?
   c. Is the respective decision made by the system appropriate in the respective context?

Niklas Schörnig

Dr. Niklas Schörnig is senior research fellow with the Peace Research Institute Frankfurt (PRIF), Germany, and visiting lecture at Goethe-University, Frankfurt. He received his Ph.D. in 2005 with a thesis on American defense industrial policy during the 1990s. In 2012 he received the "Best Article Award 2006-2011" of the German Zeitschrift für Internationale Beziehungen (Journal of International Relations). His research focuses, inter alia, on current trends in warfare, military robotics, military AI, military missions of Western democracies and Australian foreign and security policy. His most recent publications include: „ Unmanned Systems: The Robotic Revolution as a challenge for arms control". In: Reuter, Christian (2019), IT for Peace and Security, Springer: Wiesbaden, 233-256; and "Learning Unit 15: Emerging Technologies". In: EU Non-Proliferation and Disarmament Consortium eLearning Course, https://nonproliferation-elearning.eu/learningunits/emerging-technologies/ (2017, with Frank Sauer).

Take away messages and questions

(1) The military application of "artificial intelligence" and machine learning goes beyond the so-called "killer robots". It will have overarching strategic implications. The military will especially apply AI in assistance systems. The acceleration of the processing of information, decision-making and action will result in considerable instability - especially in crises when prudent action is required.

(2) Human-Machine Interaction in the Military Sector (Manned-Unmanned-Teaming; M-UMT) will be the preferred form of application for AI systems in the military sector. However, this form of interaction will decrease the role of humans in the system.

(3) Validation and verification mechanisms promise the controlled use of autonomously acting weapon systems. The main risk, however, lies in the clash between unknown systems whose interaction behaviour is not predictable.

Ahmad Sadeghi

Ahmad-Reza Sadeghi is a full Professor of Computer Science at the Technische Universität Darmstadt, where he heads the System Security Lab. Since 2012 he is also the Director of Intel Collaborative Lab for Secure Computing and Autonomous and Resilient Systems at TU Darmstadt. He is also the Speaker of the Profile Area Cybersecurity (CYSEC) of TU Darmstadt. Professor Sadeghi has been contributing to various areas of Security and Privacy research such as Trusted Computing, Mobile Security, Hardware and Software Security as well as Applied Cryptography.

Professor Sadeghi has been awarded with the renowned German prize “Karl Heinz Beckurts” for his research on Trusted and Trustworthy Computing technology and its transfer to industrial practice. Further, his group received the German IT Security Competition Award 2010. In 2018 he received the ACM SIGSAC Outstanding Contributions Award for dedicated research, education, and management leadership in the security community and for pioneering contributions in content protection, mobile security and hardware-assisted security.

<u>Take away messages and questions</u>

(1) AI is high potential threat without security: Today AI is a hype as it was in 80s, and although AI is still in its infancy it may become a sweet and bitter reality in the near future due to technological advances. However, without adequate security, privacy and accountability measures, AI has high potential to become a real threat to modern societies. Adversarial AI, which is currently a lively research area, will become an instrument of abuse in hands of professional hackers, corporates as well as nation states in future.

(2) With greater AI, comes greater responsibly: Whoever has the best AI, may also partially rule the world (control financial markets, discriminate, etc.)! While corporations are (mis)using our data to feed and improve their AI algorithms and systems, it seems that our governments, and our societies are not prepared to face the upcoming challenges that real sophisticated AI will pose on us.

(3) AI empowers corporates: Government and politicians in rich countries are providing millions of dollars funding for AI research. They run behind the trends and their own agenda. These efforts seem hasty and clueless. On the other hand, the large tech enterprises are headhunting AI experts including many academia offering them huge chunks of money and data. Consequently, the academic research on AI may play only a marginal role in a cyberworld, where data analytics are controlled by facebook, Google, and friends.

Venue
Representation of the State of Hessen to the EU
Rue Montoyer 21, 1000 Brussels

Further information available at
www.uni-frankfurt.de/science-policy or
science-policy@uni-frankfurt.de

An event by

In cooperation with

With the friendly support by

Information on the science year is available at
www.wissenschaftsjahr.de