Emerging & disruptive technologies and nuclear weapons decision making: Risks, challenges & mitigation strategies

A nuclear and new tech report

Dr Katarzyna Kubiak and Sylvia Mishra
7 December 2021
The European Leadership Network (ELN) is an independent, non-partisan, pan-European network of nearly 300 past, present and future European leaders working to provide practical real-world solutions to political and security challenges.

Dr Katarzyna Kubiak is a Senior Policy Fellow on nuclear and arms control policy at the ELN. Previously, she was a Transatlantic Post-Doc Fellow for International Relations and Security at the Norwegian Institute for Defence Studies (IFS), an associate at the German Institute for International and Security Affairs (SWP), a research assistant at the Institute for Peace Research and Security Policy (IFSH), a field researcher for the National Consortium for the Study of Terrorism and Responses to Terrorism (START) and a fellow in the German Bundestag. Following her PhD thesis on NATO nuclear extended deterrence her research areas include nuclear arms control and disarmament, nuclear non-proliferation as well as ballistic missile defense.

Sylvia Mishra is a New Tech Nuclear Officer at the European Leadership Network and a doctoral researcher at the Department of Defence Studies, King's College London (KCL). Her research focuses on nuclear strategy and nonproliferation, Southern Asian security, grand strategy and emerging technologies. She Chairs the CBRN Working Group for Women of Color Advancing Peace and Security (WCAPS) and is a N-Square Innovators Network Fellow. Previously, Sylvia was an India-US Fellow at New America, Accelerator Initiative Fellow at the Stanley Center for Peace and Security, a Scoville Fellow at the Nuclear Threat Initiative, Visiting Fellow at the Center for Nonproliferation Studies, and worked in New Delhi at the Observer Research Foundation on India-US defense and security ties.
Introduction

On 20th September 2021, The European Leadership Network (ELN) organised a young generation workshop in cooperation with the Oracle Partnership, BASIC, and the Heinrich Böll Foundation (HBS), to assess risks, challenges, and mitigation strategies for nuclear weapon decision making under technological complexity. The workshop was funded by the German Federal Foreign Office and the Heinrich Böll Foundation.

During a three-hour online workshop, participants discussed how emerging technological trends in drones, space threats, cyber threats, deep fakes, artificial intelligence (AI), and quantum technology, both individually and in tandem, will result in technological complexity for nuclear decision-making processes. The workshop used a contextualising scenario that had previously been conducted by a group of former high-level decision-makers (you can find the report for this exercise here).

Believing that the nuclear-decision makers of tomorrow could be with us today, this time we wanted to engage a younger generation of experts in a similar conversation. Our starting assumption was that this younger generation carries different expectations of emerging and disruptive technologies (EDTs) and, unburdened by past deep involvement in nuclear decision making, would offer a unique perspective on the challenges we face. Participants in this exercise consisted of nineteen early-to-mid-career participants representing academia, think tanks, and international governmental organisations. Some of these participants were drawn from the Younger-Generation Leaders Network (YGLN), which sits within the ELN, and BASIC’s Emerging Voices Network (EVN), which has a Disruptive Technologies working group.

Gender and geographical diversity of workshop participants

Below you can find reflections on some points made and discussed at the workshop. The opinions in this report represent the views of the authors, and do not necessarily reflect the position of either of the organising parties, any of its members, or the workshop participants.
Risks and challenges

- **Non-linear escalations.** Traditionally, there is a tendency to think about conflict in a linear manner (a straightforward pathway to crisis escalation). Yet, as some participants pointed out, escalation can be intentional and states may seek to manage the escalation process which traditionally does not follow a linear trajectory. Others stated that technological complexity might allow actors to skip a few steps on the escalation ladder. We might be moving away from predictable escalatory pathways; instead, crisis escalation may follow a ‘wormhole’ dynamic whereby competing states could jump between sub-conventional and strategic levels of conflict in accelerated, non-linear ways.\(^1\) The idea of non-linear pathways of escalation further gains plausibility when non-state actors are involved and attribution becomes a laborious, time consuming, and complex process, which requires a mix of technical, social, and political expertise.

- **Emerging and disruptive technologies (EDTs) might impact the 3C’s (Capabilities, Communications, Credibility) of nuclear deterrence and challenge classical deterrence.** As advancements in EDTs potentially render second-strike retaliatory forces vulnerable, they spell an uncertain future for the foundations of nuclear deterrence. Rose Gottemoeller refers to this uncertainty as the ‘standstill conundrum’, in which the lack of clarity on nuclear responses to a nuclear attack, due to loss of secure second-strike retaliatory forces, leads to escalatory pressures and instability.\(^2\) Given that EDTs may have a profound effect on second-strike capabilities, the extent of the impact on the 3C’s may also accelerate changes in military doctrines, declaratory policies, and nuclear postures. There needs to be
a deeper study on how EDTs affect the 3C’s of nuclear deterrence, thereby challenging classical deterrence and nuclear weapons decision-making processes. Studies also needs to be undertaken to quantify the uncertainties that come from incorporating EDTs in military doctrines.

• **Premature technology deployment and deployment of technologies to premature environments.** The United States Deputy Secretary of Defense, William J. Lynn III, once said “Few weapons in the history of warfare, once created, have gone unused.” Similarly, in President Harry Truman’s radio address to the American people after the atomic bomb was dropped on Hiroshima, he stated “Having found the bomb, we have used it.” The idea of the inevitability of the use of weapons proven or projected to be effective is not new. Historically, it has been the case that the production and employment of weapons (for example, the crossbow and various firearms) has proceeded despite attempts to regulate the advancing technology. Several EDTs are still in development, and their full effect, applications, and implications remain unknown. However, states might be tempted to prematurely deploy and use EDTs (that are still in the development stage) to gain first mover advantage and understand the extent of harm a specific technology can do. These can have disastrous consequences. At the same time, the premature delegation of a weapon system by one state can accelerate and encourage other states and adversaries locked in classical security dilemmas to do the same.

  “States might be tempted to prematurely deploy and use EDTs to gain first mover advantage”

• **The impact of bias and culture.** The biases involved in the nuclear decision-making process are more complex, and deeper and broader than that of the individuals involved in the decisions making process. It has personal, cultural, and technological characteristics. It starts with biases inherent to the technology developers, those making policy, decision makers, and those using the technologies in the battlefield. Technologies also have bias, depending on the data they
are fed. For example, popular voice assistants like Apple’s Siri or Amazon Alexa have been accused of displaying a gender bias. Models assessing emotion in speech display similar problems, although applying modifications have been shown to effectively mitigate these biases. Moreover, when giving technologies more room for autonomy, we may experience technology-generated bias as well. Personal biases, like worst-case thinking or process biases involving the type of information presented to the decision-maker at the beginning of the process, could also affect conclusions. Views on the usage of certain technologies can further vary depending on different government bureaucracies’ strategic cultures.

- A decision may be more susceptible to bias when made in the context of uncertainty or when the requirement of deciding with complete information cannot be met. Additionally, according to some participants, the challenge here is not just about the decision-making process or understanding how information has been collected, processed, and analyzed, but also the way a decision-maker processes this information, which is a social factor. A big challenge is thus to understand that bias is not a random error but it is structurally inherent in any process that involves human decision making. Finding ways to enable decision makers to acknowledge their own biases and minimise them to the greatest extent possible will therefore be key. It is thus unsurprising that the latest NATO AI Strategy discusses the need to minimise unintended biases, consciously bridging biases, and promoting bias mitigation practices in use.

- **Lack of awareness of being in or out of the loop.** Several participants voiced concerns about decision makers’ understanding of whether they would be “in” or “out of the loop” in each individual instance of a decision-making process and how that affects the question of human agency in nuclear weapon use decision-making. Additionally, it is not just about whether decision-makers understand their role in the process but also more objectively whether the system itself provides the opportunity to exercise human judgement and control.

- **Trust in advice.** It is unlikely that decision-makers will have the capacity to internalise all information available, hence they will inevitably need to trust their advisors. As such, many of the challenges a decision-maker will face are spread across various stages and actors. At the same time, levels of confidence in intelligence, political advice, and direction might not be uniform. For instance, decision-makers might trust intelligence and information received differently, and in accordance with her own cognitive biases. This could generate stress, especially in newly elected decision-makers. Drawing from the literature on ‘anchoring bias psychology’, participants also pointed out that decision makers may rely heavily on the first piece of information they received, which could chart the course of processes in the decision-making chain.
• **Risk manipulation.** Thomas Schelling once wrote that “If brinkmanship means anything, it means manipulating the shaved risk of war. It means exploiting the danger that somebody may inadvertently go over the brink, dragging the other with him.” Participants mentioned that some actors may use EDTs to exploit the risk of escalation and gain coercive leverage. North Korea’s pursuit of nuclear weapons is an example of ‘signature brinkmanship’ - a strategy that Pyongyang uses to bring other states to the negotiating table. As both nuclear and non-nuclear weapons states expand their EDT inventories, several states and non-state actors might utilise risk manipulation and engage in brinkmanship to create tensions, instability, and gain an upper hand in negotiations.

> “Several states and non-state actors might utilise risk manipulation and engage in brinkmanship to create tensions, instability, and gain an upper hand in negotiations.”

• **Domestic communication, public pressure, and audience cost.** New channels of communication have amplified the pace at which information travels and democratised ways to curate it. Social media has the potential to exacerbate tensions during crises, with the general population falling prey to disinformation campaigns and provocations. The Covid19 pandemic showcased the rapid pace at which false information, misinformation, and disinformation travels and the rapid pace at which malicious actors can respond to - and leverage - emerging (and emergency) situations. Participants discussed how the multiplicity of information platforms during crisis and peacetime may affect domestic audiences. Will a policymaker be able to communicate effectively to a domestic audience, convey message to an international audience, and signal to adversaries? What impact will they have? Can governments utilise social media platforms effectively for political mobilisation? Will the public take to social media platforms to pressure their government to take actions? Will public pressure complicate the decision-making atmosphere?

• **Credibility of information.** In the background of brittle trust among states and even among multiple actors within states, it is increasingly challenging to receive and verify correct information. Given the spectrum of untrustworthy information,
the chances of erosion of trust in decision makers’ own systems has grown. For instance, there doesn’t need to be a convincing deep fake, just the threat of a convincing deep fake (i.e., to raise doubts in one’s own intelligence which could lead to instability). How can decision-makers thus mitigate risks from overreliance on limited information, unreliable, or representative information? How can we mitigate the risk of harm from the use of inaccurate or false information?

- **The challenge of attribution.** With the introduction of a slew of EDTs and their use in conflicts, one of the immediate challenges for states is determining attribution. The lack of swift attribution (as was seen in the case of coordinated drone and cruise missile strikes on Saudi oil facilities in 2019) might impede a state’s ability to undertake immediate retaliatory measures. It may even give rise to misplaced judgement due to incorrect or incomplete information. These can impact strategic stability, thereby lowering thresholds for strategic patience.

“\[The lack of swift attribution might impede a state’s ability to undertake immediate retaliatory measures or give rise to misplaced judgement\]”

- **An acute lack of understanding or awareness of adversaries’ EDTs capabilities and decision-making processes.** Traditionally, in a moment of crisis, a decision-maker would likely need to know two sets of information. First, what capabilities does the adversary have, what can and might they do with them, and how can we stop them? Second, what are my own capabilities and what political and military goals and objectives can I achieve with these? Additionally, in times of technological advancements, the decision-maker may want to know about the adversary’s relationship with specific technologies, their development process, or how they train to use a technology. Multiple EDTs are improving the situational awareness by offering better and refined data about adversaries’ capabilities. Such awareness can have a stabilising effect. However, because EDTs are still new and often invisible (like AI capability or cyber capability) the decision-maker might experience lack of understanding and awareness of adversaries’ related
capabilities and decision-making processes. This creates further challenges. For example, from the standpoint of nuclear weapons, decision makers from China and Pakistan (countries who are “lagging” in terms of conventional or strategic capabilities) who are trying to bridge their capability gap vis-a-vis primary adversaries like the US and India, respectively, will find less to little incentives for risk mitigation. In fact, EDTs will bridge capability gaps vis-a-vis adversaries and augment ambiguity, thereby bolstering deterrence. Additionally, as EDTs evolve, it might be challenging to have full confidence in one’s own capabilities and battle-field readiness.

• **Miscalculations, inadvertent escalations, and unintended consequences.** Participants mentioned that the development and deployment of EDTs raise the chances of incidents, accidental war, miscalculations, and accelerate inadvertent escalation. Workshop participants are not alone in this judgement. Research reveals that many experts expect EDTs to be destabilising and increase the chances of accidental crisis escalation. We have seen such close calls before; for example, in 1983, the Soviet officer Stanislav Petrov judged reports from the early warning system as faulty and, by disobeying orders, did not escalate the incident (which indeed turned out to be a false alarm) higher up the decision ladder, thereby possibly averting thermonuclear war. Another example was the 1995 Norwegian missile incident. In that instance, Russian systems initially read a US and Norwegian missile carrying scientific equipment as a possible US attack on Russia. Russian President Boris Yeltsin was faced with the option to pre-emptively retaliating, which he declined. EDTs like AI furthermore increase the risk of escalation due to their complex and opaque functioning. In this context, operators of today’s advanced AI-based systems might understand even less than operators of the less automated/advanced systems of the Cold War, in regards to how these systems function and how to identify errors in a crisis.
Mitigation strategies

The Young Generation Workshop generated a highly engaging discussion. It underscored that the young professionals of today, who will become the leaders of tomorrow, are deeply and purposefully thinking about the changes and complexities posed by emerging technologies in the nuclear policy space. Participants also reflected on several ideas on how to start mitigating related risks:

- **Understand the subject better.** The nuclear community at large needs more structured and systematic efforts in assessing technological trends, distinguishing hype from facts (e.g., by using the Gartner hype cycle), and quantifying uncertainties.

- **Clarify intent.** In a complex and multiplex deterrence relationship era, it is important that countries clearly signal their intent through declaratory policies and red lines. Defining what is ‘acceptable’ and ‘unacceptable’ requires the buy-in of all parties involved. While the United States and the United Kingdom have adopted declaratory policies at high-level, other nuclear-armed states also need to review, update, and share their strategic thinking about new military applications and uses of technologies. As different individual technologies are at different stages of development, continuous review processes will be essential. These could also pave pathways for multinational cooperation on developing new treaties and codes of conducts on the usage of EDTs. For instance, nuclear armed states could come to an understanding and a binding agreement of non-attack on nuclear
command and control infrastructure. Non-Nuclear Weapons States (NNWS) with high cyber offensive and defensive capabilities could also attack nuclear command, control, and communications (NC3) of nuclear weapons states.

“Nuclear-armed states need to review, update, and share their strategic thinking about new military applications and uses EDTs.”

- **Keep communication lines open and have off-ramps to correct misinformation and misunderstandings.** Establishing a widespread communication infrastructure, direct lines, and channels between adversaries (including through trusted allies) is necessary to offer opportunities for clarification when concerns regarding specific activities arise. Setting such lines up during crisis or war might not be feasible. It is thus always important to keep communication lines open, regardless of the nature of the relationship between states. Elongating decision-making time can buy clarification time that may be needed for attribution or communication with an adversary.

- **Reduce biases.** More work needs to be done to understand and prevent biases at all responsibility levels of nuclear decision-making (technological bias, technology development bias, bureaucratic biases, advisory and advisory process bias, and nuclear decision-maker biases). When integrating new technologies into nuclear decision-making, actors should be cognizant of existing and ongoing work on biases in sociotechnical systems.

- **Do not assume everybody wants to reduce risks.** Those who have an interest in exploiting risks of escalation to gain coercive leverage, including through EDT's, may not seek to reduce risks.

- **Train a diverse cadre of nuclear decision-makers, mid-career professionals, and leaders of tomorrow.** Leaders of tomorrow need to be creative thinkers and problem solvers from diverse backgrounds. They need training to develop functional understandings of technologies, their full potential, limitations, and risks. Future decision-makers should undergo exercises, war games, and other simulations to be able to apply policy guidelines and strategies, and to understand what they
can and cannot do under international and national laws. Studies to determine baseline technological literacy of decision-makers could also give suggestions on where education and exercise is needed.

“Leaders of tomorrow need to be creative thinkers and problem solvers from diverse backgrounds.”

- **Encourage private sector to understand the security implications of their innovations.** Several private sectors companies, particularly digital tech companies, are already involved in promoting peaceful use of Information and Communications Technologies (ICTs) or EDTs. However, more needs to be done to integrate tech companies and other private sector stakeholders into ongoing and future diplomatic discussions that ensures EDTs are developed and managed in a peaceful and stable manner to reduce nuclear risk associated with EDTs. Unless all levels of the private sector fully comprehend the security implications of the technologies they are developing a gap will persist between the original intention behind a specific technology and its end-use. This has already had real life consequences on several occasions. In 2018, for instance, several employees quit over Google's involvement in an artificial intelligence drone program for the Pentagon called Project Maven. Nearly 4,000 workers petitioned and demanded to end Google's participation in it and eventually pushing Google to not renew its contract for Project Maven.\(^{14}\)

“Unless all levels of the private sector fully comprehend the security implications of the technologies, they are developing, a gap will persist between the original intention behind a specific technology and its end-use.”
• **Community integration.** In the backdrop of an emerging ecosystem of multiple stakeholders in EDTs, there is an urgent need to de-silo and integrate relevant communities, promote collaboration, and develop joint and cohesive mitigation strategies. This includes establishing dialogue between agencies responsible for different domains to understand interlinkages. Multilateral approaches should also aspire to transfer collective knowledge from different industry branches to governments and encourage exchanges between the countries that have a very developed understanding of the intersection of nuclear weapons and EDTs and those that do not.

• **Use existing structures to design risk mitigation strategies.** While state actors have not yet begun to jointly address the challenges of new technologies for nuclear decision-making, different expert communities, such as Computer Security and Incident Response Teams (CSIRTs)\(^ {15}\) or the Forum of Incident Response and Security Teams\(^ {16}\), already respond and solve day-to-day cybersecurity problems through diverse national, regional, and international networks. As such, reinventing the wheel is not necessarily needed in looking for collaborative solutions to mitigate risks. Instead, we can build on what already exists.

> “Reinventing the wheel is not necessarily needed in looking for collaborative solutions to mitigate risks. Instead, we can build on what already exists.”

The ideas offered and discussed during the ‘Young Generation Workshop’ on EDTs and nuclear weapons decision-making had several convergences and similarities with that of the ideas exchanged during ELN’s January 2021 workshop with former high-level nuclear decision-makers and current officials.\(^ {17}\)

Both established experts and emerging professionals agreed that EDTs and their combined effect will have an impact on decision-makers’ ability to manage, assimilate, interpret, trust, verify information, and, ultimately, make a nuclear decision. There was broad agreement that different EDTs working simultaneously would complicate the decision-making processes, especially amidst shortening decision-making timelines.
Both the January and September 2021 workshops hosted by the ELN and its partner organisations showcased that EDTs pose challenges and offer opportunities to the existing global nuclear order. Both workshops - and especially the September workshop with younger participants - underscored that whilst EDTs complicate nuclear weapons policy making, there are means and measures by which EDTs can be better managed and developed in a manner that is stabilising. The emerging professionals also emphasized the importance of bridging the gap between policymakers and technocrats, and training professionals with both policy and tech backgrounds. Overall, the workshop highlighted the importance of undertaking focused and analytical research and studies on EDTs and their impact on nuclear weapons policies.
Endnotes


Annex: List of participants and organisers

List of participants

Julia Balm, PhD Candidate, Freeman Air and Space Institute, King’s College London (Canada)
Belen Bianco, UNIDIR (Argentina)
Julia Cournoyer, Research Assistant, International Security Programme, Chatham House (UK)
Anuradha Damale, Research Assistant, VERTIC (UK)
Marina Favaro, Visiting Research Fellow, IFSH (Canada)
Niamh Healy, PhD Researcher, University College London (UK)
Verena Jackson, Center for Intelligence and Security Studies (Germany)
Patricia Jaworek, Consultant, Nuclear Threat Initiative (Germany)
Raquel Jorge-Ricart, Berkman Klein, Center for Internet & Society, Harvard University (US)
Rafael Loss, Coordinator for Pan-European Data Projects, European Council on Foreign Relations (Germany)
Christoph Mayer, Tech Governance Research Fellow at the German Council on Foreign Relations (Germany)
Farzan Sabet, Researcher, UNIDIR (Iran)
Jantje Silomon, Researcher, IFSH (Germany)
Dmitry Stefanovich, Russian International Affairs Council (RIAC) Expert and a non-resident Fellow with IFSH Hamburg and Research Fellow at the Center for International Security, Primakov Institute of World Economy and International Relations (IMEMO RAS) (Russia)
Lydia Wachs, Research Assistant, Stiftung Wissenschaft & Politik (Germany)

Organisers

Julia Berghofer, Policy Fellow and Project Manager for the Younger Generation Leaders Network (YGLN) (Germany)
Emily Enright, Policy Fellow, BASIC (UK)
Giorgio Franceschini, Head of Foreign and Security Policy Division, Heinrich-Böll-Stiftung (hbs) (Germany)
Milena Grünwald, Project Officer, Foreign and Security Policy Division, Heinrich-Böll-Stiftung (hbs) (Germany)
Peter Kingsley, Chairman and Co-founder, Oracle Partnership (UK)
Katarzyna Kubiak, Senior Policy Fellow, European Leadership Network (ELN) (Poland/Germany)
Marion Messmer, Co-Director, BASIC (UK)
Sylvia Mishra, New Tech & Nuclear Officer, European Leadership Network (ELN) (India)
Chris Spedding, Policy Fellow, BASIC (UK)
Graham Stacey, Senior Consulting Fellow, European Leadership Network (ELN); Former Chief of Staff of NATO Transformation (UK)
Simon Tilford, Director, The Oracle Partnership (UK)
Sebastian Brixey-Williams, Co-Director, BASIC (UK)
**The European Leadership Network**

The European Leadership Network (ELN) is an independent, non-partisan, pan-European network of nearly 300 past, present and future European leaders working to provide practical real-world solutions to political and security challenges. The ELN builds better security for wider Europe through its research, publications, events, practical policy advocacy, media reach and high-level networks. It concentrates on what it judges to be the gravest risks to Europe’s security and on the risks where it assesses that it can make the greatest difference.

**The Oracle Partnership**

The Oracle Partnership brings together some of the world’s leading domain experts, well-proven foresight, scenario and strategy methodologies and a range of state-of-the-art artificial intelligence (AI) tools, focusing on strategic risk and innovation. It develops intelligence beyond conventional futures research, looking for tell-tale early signs of political and economic disruption, policy changes, sudden shifts in public sentiment and breakthrough technologies, long before they go mainstream. The goal is to model complexity and uncertainty, creating strategic frameworks for organisations to navigate emerging reality.

**The British American Security Information Council**

The British American Security Information Council (BASIC) is a non-partisan think tank based in London. It takes an inclusive approach to promote nuclear disarmament and non-proliferation by working with politicians, civil society, and other people who share this vision, as well as with those who might oppose it. BASIC promotes meaningful dialogue amongst governments and experts in order to build international trust, reduce nuclear risks, and advance disarmament. It envisions a world that uses cooperative measures, rather than the threat or use of force, to achieve peace and security. BASIC has a global reputation for convening distinctive and empathic dialogues that help states overcome complex strategic and political differences.

**The Heinrich Böll Foundation**

The Heinrich Böll Foundation (hbs) is a political foundation close to the party Bündnis 90 / Die Grünen. The foundation sees itself as an agency for green ideas and projects, as a reform policy future workshop and international network with partner projects in around 60 countries. Heinrich Böll’s encouragement of civil society interference in politics is a model for the work of the foundation. Its primary task is political education in Germany and abroad to promote the democratic will, the socio-political commitment and international understanding. It is guided by the basic political values of ecology, democracy, solidarity, and nonviolence.