Wilton Park

Report

Emerging technologies and the Nuclear Non-Proliferation Treaty

Monday 26 - Wednesday 28 September 2022 | WP3091

Wilton Park

Report

Emerging technologies and the Nuclear Non-Proliferation Treaty

Monday 26 - Wednesday 28 September 2022 | WP3091

In partnership with the Stanley Center for Peace and Security, and King's College London with support from the MacArthur Foundation

The impact of emerging technologies has assumed greater saliency in recent years, but their implications for nuclear risk and the global nuclear order remain contested and uncertain. Applications of some technologies could have the potential to exacerbate nuclear risks. There may also be ways to leverage these technologies in support of non-proliferation, disarmament, and arms control efforts.

This Wilton Park conference sought to initiate an international dialogue to better understand the role of emerging technology and nuclear risk. In particular, it aimed to:

- Develop understanding and awareness of the roles that emerging technologies play in the global nuclear order.
- Identify risks and opportunities associated with emerging technologies and the implications for crisis stability, arms control, and the NPT.
- Bridge gaps in thinking between policy and technical communities.
- Define a forward-looking agenda for emerging technology within the NPT.

Introduction

1. Policymakers are facing an increasingly complex and competitive technological and information environment and it is critical to better understand emerging technologies, their impacts on international peace and security, and prospects for regulations and arms control. As such, there is a heightened demand for both policy and academic communities to work on emerging technology issues.
2. To this end, Wilton Park convened a group of experts to explore the following questions:
    - What is the role of emerging technologies in international peace and security?
    - What are the risks and benefits of emerging technologies for arms control?
    - What is the relationship between political and technical communities? What are the overlaps and differences between thinking in policy vs. technical communities?
    - What does all of this mean for the NPT?
3. Several themes garnered consensus throughout the conference, including:

- The importance of defining key concepts and making issues “digestible” in order to facilitate useful conversations and progress on these topics.
- The importance of assessing both the risks and opportunities created by emerging technologies. Of note, emerging technologies present challenges and opportunities to geopolitical cooperation and competition, crisis stability, and the monitoring and verification of arms control agreements.
- The importance of fora in which conversations of emerging technology take place, as well as the actors involved in those conversations.

Definitions

4. The Institute for Peace Research and Security Policy (IFSH) at the University of Hamburg provides a useful definition of emerging technologies as "those technologies, scientific discoveries, and technological applications that have not yet reached maturity or are not widely in use but are anticipated to have a major – perhaps disruptive – effect on international peace and security." [1]
5. Several problems, however, underlie the term “emerging technologies,” particularly since it depends on human perceptions of what is emerging. The threshold of when a technology can be considered mature is subjective and dependent on the criteria by which individual scholars evaluate maturity against. Technological maturity can be assessed in a number of ways, including through the development of the technology itself, the implementation of its applications or uses, and the discovery of its intended and unintended impacts. A technology considered mature in terms of development and applications may later develop new applications and impacts, potentially necessitating its reclassification as an emerging technology. Similarly, new technologies may interact with existing technologies in ways that alter its impacts. Given that a technology's status as emerging or mature is not static, an iterative approach is necessary when classifying emerging technologies.
6. While IFSH’s definition of emerging technologies is not immune to these definitional issues, it is sufficient enough to enable policymakers to address the impact of these technologies before the window to do so has passed, while simultaneously avoiding exaggeration and alarmism. Ultimately, policymakers and scholars must take a pragmatic approach to defining and operationalizing the term "emerging technologies" and avoid definitions that are too broad or too narrow.

Emerging technologies and implications for the global nuclear order

Crisis Stability

7. Emerging technologies have varying impacts on crisis stability, which is the absence of incentives to use nuclear weapons first in a crisis. Crisis stability is not an absence of conflict, but rather the ability to put a cap on escalating levels of violence. When escalation becomes more uncontrollable, crisis stability worsens as pressure to initiate a first nuclear strike builds. Emerging technologies, in particular, can alter the perceptions held by leaders and the general public. Perceptions underpin crisis stability, especially perceptions about the security of a state’s second-strike capability. The status of the geopolitical environment—particularly whether it is peacetime or a time of conflict—is also important to how actions are perceived, and therefore how escalatory actions are.

[1] Marina Favaro, Neil Renic, and Ulrich Kühn. “Negative Multiplicity: Forecasting the Future Impact of Emerging Technologies on International Stability and Human Security.” Institute for Peace Research and Security Policy at the University of Hamburg, September 26, 2022. https://ifsh.de/en/news-detail/new-research-report-forecasting-the-future-impact-of-emerging-technologies-on-international-stability-and-human-security. Pg. 16-17.

8. One key element of crisis stability is a leader's decision space, defined by the availability of a strategy set and the time to process crisis data in an information rich environment. Emerging technologies have varying effects on this space. In some cases, they may reduce the available processing time or availability of response options, whereas other technologies may increase these. For instance, AI enabled situational awareness may detect missile launches and relay that information across a network faster, providing leadership more time to plan out a response.

9. As more actors acquire emerging technologies, states lose control over disclosure decisions, "the choice governments make regarding whether, when, and how to release sensitive information about foreign actors.[2] More accessible commercial satellite imagery is a clear example of this, as private actors can observe and disclose certain state actions that would otherwise be secret. In altering a state's control over disclosure decisions, emerging technologies can limit the strategy set or response options critical to a leader's decision space.

10. States continue to incorporate emerging technologies into their militaries, in turn creating challenges for crisis stability by impacting policymakers' perceptions of adversaries' actions and intentions and diminishing confidence in strategic systems. First, state manipulation of emerging technologies can exacerbate ambiguity about adversaries' actions and intentions. AI-enhanced deep fakes, for example, can create doubt in policymakers' minds about adversarial actions as they appear to occur, and therefore the accuracy of the information on which they base their decisions. Moreover, AI can cause policymakers to doubt their interpretations of adversaries' intentions. One participant noted that when an adversary uses AI-enabled C3 systems, for example, policymakers may struggle to interpret the intentionality behind actions, and therefore will not know how to respond.

11. The greatest risk emerging technologies pose to crisis stability lies in their ability to diminish confidence in strategic systems by endangering—either kinetically or non-kinetically—the functionality of the systems. Strategic systems underpin crisis stability by enabling a secure second-strike capability. In threatening these assets and undermining perceptions of a secure second-strike, emerging technologies intensify pressures of intentional and unintentional escalation. This impact is especially salient when technologies can threaten assets of nuclear command, control, and communications (NC3) with a high degree of conventional-nuclear integration (CNI).

12. ASATs, in particular, can directly target space-based NC3 assets, or otherwise endanger NC3 with debris from the destruction of nearby space objects. Additionally, the risks cyber operations pose to strategic assets will intensify in the future as states operate in increasingly networked environments with high levels of CNI. In these conditions, cyber operations, particularly those that can deny, degrade, or destroy NC3 will have a highly escalatory effect. The escalatory effect of cyber operations, however, may be unintentional in some cases. For example, cyber espionage in NC3 satellites can be misinterpreted as an attempt to disrupt NC3 altogether, creating the perception of an insecure second-strike capability and subsequently triggering escalation.[3]

[2] Lin-Greenberg, Erik, and Theo Milonopoulos. "Private Eyes in the Sky: Emerging Technology and the Political Consequences of Eroding Government Secrecy." Journal of Conflict Resolution 65, no. 6 (2021): 1067–97. https://doi.org/10.1177/0022002720987285. Pg. 1068.

[3] James M. Acton; Escalation through Entanglement: How the Vulnerability of Command-and-Control Systems Raises the Risks of an Inadvertent Nuclear War. International Security 2018; 43 (1): 56–99. doi: https://doi.org/10.1162/isec_a_00320. Pg. 61.

13. Emerging technologies also create several opportunities for crisis stability, specifically in terms of hardening systems and strengthening intelligence, surveillance, and reconnaissance (ISR) capabilities. Emerging technologies can harden and secure systems, which will be especially important as states roll out and maintain highly networked systems for their militaries featuring CNI. By hardening strategic systems, for example with quantum technology, states decrease non-kinetic vulnerabilities to NC3 and secure second-strike capabilities, in turn reinforcing crisis stability. Cyber capabilities and AI can also secure systems through continuous network monitoring and autonomous patching.

14. Emerging technologies can also strengthen crisis stability through advanced ISR capabilities. Leveraging emerging technologies to provide policymakers with unprecedented levels of information and situational awareness can, in theory, alleviate some of the ambiguity that creates escalatory pressures undermining crisis stability. Advanced sensor technology and social media, for example, can provide clearer or entirely new views into the strategic and operational environment in real-time. Emerging technologies can also help intelligence analysts and policymakers understand the information environment more efficiently. In particular, AI and machine learning can provide an initial scrub of intelligence, supporting analysts as they sort signals from noise.

Arms Control and Disarmament

15. Perhaps the most widely discussed area of emerging technologies' impact on the nuclear policy landscape is the use of emerging technologies to verify and monitor arms control and disarmament agreements. The use of satellite imagery by open source intelligence (OSINT) analysts to assess compliance with arms control has gained widespread notice by the media and policymakers alike.

16. Widespread access to satellite imagery from private companies is used to monitor North Korean ballistic missile tests, Russian nuclear modernization, Iranian construction at nuclear and missile sites, and military troop deployments during times of military conflict. More recently, researchers have been able to access more advanced forms of remote sensing such as thermal, hyperspectral, and synthetic aperture radar imaging. These all may provide unique benefits. For instance, researchers at 38 North have used thermal imagery in an attempt to assess when the Yongbyon reactor in North Korea was producing plutonium.[4]

17. While satellite imagery catches most of the media attention, the tools available to civil society for monitoring and verification span a broad range of technologies ranging from analysing complex seismic and space-based sensor data to more simple monitoring of things such as social media posts. The 21st century is full of billions of sensors that are all collecting data. Societal verification's potential to collect data spans open-source communities' ability to analyse data that comes from cell phones, computers, trade databases, seismic sensor data, and ship and airplane tracking data. The confluence of many emerging technologies being used to develop multi-source verification has greatly enhanced civil society's ability to participate in arms control compliance and verification discussions.

18. Late 20th and early 21st century technological innovations have greatly expanded the amount of data collection methods that societal verification has access to. One emerging problem that plagues virtually every field, including open-source intelligence, is what to do with all of this data. Civil society is only starting to utilize new tools to interpret and visualize the data that has been collected. One emerging answer to the visualization challenge has been an interest in enabling technologies such as machine learning and artificial intelligence.

[4] Bermudez, Joseph et al., Youngbyon Facility: Probable Production of Additional Plutonium for Nuclear Weapons, 38 North, July 2017.

19. Researchers, governments, and international organizations have already begun using machine learning tools, such as text-based analyses, to monitor WMD related activities. For instance, the IAEA has created a Content Reification Engine, called ICORE, that examines open-source reporting to identify possible undeclared nuclear fuel cycle activities.[5] The use of machine learning to mine or sift through large databases for keywords can be utilized by civil society to monitor scientific publications coming out of countries such as Iran or North Korea. These sorts of applications are fairly feasible by many organizations working in societal verification today.

20. Other discussions seek to use artificial intelligence to accomplish much more complex tasks such as image classifications. Some researchers envision the use of artificial intelligence to scan large databases of commercially purchased satellite imagery and recognize objects, including missiles.[6] Other researchers have contemplated the use of deep neural learning to identify the sale of proliferation sensitive or export-controlled technologies and bolster strategic trade controls.[7] These tasks are riddled with logistical challenges today, but may be consistently viable in the future.

21. Emerging technologies may provide unique opportunities for arms control and disarmament efforts by developing novel approaches to verification in the public sphere and by bolstering the capabilities that governments have at their disposal to verify nuclear arms reductions.

22. OSINT provides the public with many unique inputs into discussions of verification and compliance of nuclear arms control. More eyes watching arms control agreements may mean that it is easier to detect and verify defection. Moreover, publicly available data that demonstrates non-compliance can help bolster confidence in claims of defection. This increased transparency might make monitoring and verification of future arms control agreements easier. These benefits would also be applicable to disarmament efforts. It is unlikely that major arms control and disarmament efforts will rely on OSINT alone for verification. However, OSINT may bolster verification arrangements of multilateral or bilateral agreements or codes of conduct. Additionally, some states may find that they can voluntarily self-report compliance with existing arms control agreements by packaging and re-releasing publicly available data to enhance confidence.

23. Key emerging technologies may also make efforts towards nuclear disarmament more verifiable. The QUAD Nuclear Verification Partnership has made some progress on using emerging technologies to process trace nuclear materials and dummy warheads, authenticate and certify of dummy warheads, and use digital data cryptography to track workflows. A remaining central challenge with verifying the dismantlement of a nuclear warhead is the need to prohibit access to information about warhead design and the fissile material inside of the warhead.

[5] Burr, Patrick et al, Using Machine Learning and Natural Language Processing to Enhance Uranium Mining and Milling Safeguards, IAEA Symposium on International Safeguards, November 2018.

[6] Hanham, Melissa and Lewis, Jeffrey, Remote Sensing Analysis for Arms Control and Disarmament Verification in Federation of American Scientists Nuclear Verification Capabilities Task Force, September 2017

[7] Withorne, Jamie, Machine Learning Applications in Non-Proliferation, James Martin Center for Nonproliferation Studies, August 2020.

24. While some emerging technologies may bolster attempts at arms control by providing enhanced transparency, other technologies may present several challenges to future arms control agreements. Novel types of uranium enrichment such as laser separation may make proliferation less visible and more difficult to detect. The use of additive manufacturing could bypass existing export control regulations. Advanced computational modelling capabilities may allow vertical proliferation without the need for a resumption of nuclear weapons testing. These technologies might also create challenges to disarmament by lowering the technical and material barriers to rearmament, creating hedging pressures and non-proliferation challenges.

25. There may also be a dark side to greater civil society involvement in the nuclear domain. Many arms control agreements are intricate, and big data may erode trust by painting a complicated picture of compliance. Deep fakes, manipulated imagery, and social media campaigns could seek to undermine faith in arms control. Additionally, OSINT analysts may simply get conclusions wrong while simultaneously increasing public pressure on political leaders to take action.

Conclusions and recommendations

26. The international community is slow to capture emerging technologies in treaties and arms control agreements. Controls often follow technology plateauing after initial development and reaching some degree of “maturity” in its applications. Preserving the benefits of emerging technologies while minimizing risks is of critical importance to international peace and security, but will be difficult to accomplish as states continue to compete in technological development against the backdrop of a worsening geopolitical environment.

27. To date, the most effective conversations about emerging technologies and nuclear risks occur outside the NPT, specifically within Track 1.5 and Track 2 dialogues, the “deterrence community,” and the P5. There was no strong consensus amongst conference participants about which fora is the most appropriate, or would yield the greatest success, in addressing these issues. Still, participants made several recommendations regarding potential regulatory frameworks, actors that can contribute to emerging technology conversations, and what progress on emerging technology issues might look like within and outside of the NPT. Even if the desired outcome is unattainable in one or all fora, the process of the dialogue itself could be valuable for fostering a shared understanding of the interaction of emerging technologies and nuclear weapons.

General

28. Participants identified two major frameworks to approach emerging technologies within or across different fora: the responsible behaviours approach and the legally-binding measures approach. With the responsible behaviours approach, states and commercial actors would theoretically act within a norms-based code of conduct, consisting of responsible innovation standards, peaceful use norms, and iterative assessments of the unintended effects of technologies. Alternatively, the legally-binding measures approach would resemble past legal frameworks that formally regulate other disruptive technologies such as nuclear weapons or bioweapons. These approaches, however, are not mutually exclusive and could be used in tandem. For instance, a responsible behaviours approach adopted by a group of states with support from commercial actors could lay the normative groundwork for more formal regulations under the legally-binding measures approach.

29. The dual-use nature of emerging technologies creates a dilemma that states must grapple with as they seek measures to maximize benefits and minimize risks—it will be increasingly difficult to distinguish between commercial, military, and other applications of emerging technologies. This is not unprecedented, though, as states historically navigated this dual-use dilemma for other technologies. For example, nuclear technology yielded not only a weapon of mass destruction but also significant opportunities for peaceful energy uses. The NPT was part of the solution to this dilemma, seeking to support peaceful uses of nuclear technologies while preventing the proliferation of nuclear weapons. Some conference participants, however, expect the line between military and commercial applications of emerging technologies to be further blurred in the future as their uses become less attributable and their development more accessible.

30. States that develop and possess emerging technologies—commonly the wealthiest or most powerful actors in the international community—have an outsized voice in conversations about emerging technologies. The impacts of emerging technologies, however, reach far beyond this group of powerful states. Conversations about responsible technological development, applications, and regulations should include a broader portion of the international community. In particular, this will be important as emerging technologies with strategic or otherwise disruptive effects become more accessible—proliferating across the world and introducing new actors into existing escalation dynamics.

31. An interdisciplinary approach to conversations about emerging technologies could also yield significant benefits by introducing new perspectives and informed analyses. Conference participants noted, in particular, that individuals from the insurance, systems engineering, and neuroscience sectors could make meaningful contributions to these conversations. For example, insurance industry experts could lend important insights into risk foresight and quantification. The neuroscience sector could also provide insights into the science behind perceptions and decision-making, which are altered by emerging technologies.

32. In order to engage different communities on issues of emerging technologies and nuclear risks, conference participants noted that bridge-building efforts will likely need to occur, in particular, within:

    - The NPT process,
    - The P5, and
    - Technical communities and private sector stakeholders.

NPT Process

33. While the 2022 NPT Review Conference did not achieve a consensus final document, there was agreement that NPT member states would put together a working group to strengthen the NPT review process. Participants at the Wilton Park conference were divided about whether this NPT working group should engage with emerging technologies directly. Despite these differences of approach, all Wilton Park participants recognized that emerging technologies will continue to be of pivotal importance and should be incorporated in more detailed discussions.

34. Conversations about emerging technologies and nuclear risks could fit well into risk reduction conversations within the NPT. However, there is a high degree of scepticism amongst NNWS about the intentions behind NWS' focus on risk reduction efforts. NNWS see broader conversations about risk reduction as a distraction from sluggish progress on disarmament. At the 2022 RevCon, for example, risk reduction was held hostage by certain NNWS looking for more substantive nuclear disarmament concessions. Still, the NPT could be a good starting for global conversations about emerging technologies and nuclear risks, raising the questions:

- What are states' (both NWS and NNWS) concerns about emerging technology and nuclear weapons?
- Does the international community have the tools, within the NPT and the review process, to address these issues?

35. There is a robust debate around whether emerging tech should be built into the structure of the NPT review process directly. This is the case for many other international arms control and non-proliferation regimes. For instance, Article XII of the Biological and Toxic Weapons Convention (BWC) stipulates that the BWC review process "shall take into account any new scientific and technological developments relevant to the Convention." A similar mandate could be included within future NPT Preparatory Committee meetings to directly engage with scientific and technological developments.

36. On the other hand, the working group tasked with strengthening the NPT review process has a very full agenda. A lack of consensus in the 2015 and 2022 NPT Review Conferences has left few states enthusiastic about the health of the NPT regime and adding additional scopes of work may further problematize the review process. Some Wilton Park participants suggested that NPT diplomats and scientific and technical advisors should meet outside of the formal NPT PrepCom meetings to discuss the impact that emerging technologies will have on the treaty.

P5 States

37. While conference participants were unsure about how effective emerging technology efforts could be within the NPT, other fora may provide better initial progress. The P5 process, in particular, could serve as an initial multilateral setting to evaluate emerging technologies' impacts and the path forward for controlling them. Emerging technologies would logically fit within existing risk reduction conversations amongst the P5.

38. The formal P5 process provides P5 states space to discuss NPT implementation and engage in confidence-building. While the P5 process is not necessarily a space for negotiations, it can serve as a venue to construct building blocks for future negotiations. Conference participants identified two initial recommendations to be concurrently pursued. First, P5 members can initiate joint studies to create mutually shared understandings of how emerging technologies may impact arms control, risk reduction, nuclear deterrence, and nuclear disarmament. These studies will help create mutually defined concepts and more digestible issue sets while providing a general understanding of nuclear risks critical to identifying the right approach to controlling emerging technologies.

39. Second, P5 states can invest in capacity-building efforts to expand the conversation. It will be critical to engage the right experts in these conversations, including individuals from the military, defence industry, interdisciplinary academic community, and other relevant fora. Moreover, the P5 process could benefit from incorporating the perspectives of emerging scholars and practitioners, such as those in the P5's Young Professional Network, so as to gain fresh insights and make capacity-building efforts more sustainable across generations.

40. The broader international community and NNWS, in particular, might perceive that P5 attention to emerging technology issues is a means for avoiding engagement in other multilateral settings, or distracting from disarmament commitments. Still, the P5 could make substantive progress on dialogues about emerging technologies, escalation pathways, and nuclear risks while attempting to respond to these concerns. For example, P5 members could issue a joint statement about how regulatory measures for emerging technologies could strengthen nuclear disarmament efforts.

Technical Communities and Private Sector Stakeholders

41. Conference participants repeatedly stressed the importance of expanding conversations on emerging technologies to include broader cohorts of individuals across different communities. In particular, dialogues should include a more diverse set of actors from the technical community and private sector.

42. The commercial sector typically leads the research and development (R&D) of emerging technologies but should be more involved in conversations about the responsibilities and intentions behind R&D. From the point of initial R&D onward, the potential first and second-order impacts of the riskiest technologies on international peace and security should be evaluated, not only by international security practitioners but also by the commercial sector. Technical and policy communities should collaborate to explore “red lines” for applications of emerging technologies.

43. While states can incentivize commercial sector involvement in broader conversations about emerging technologies, it will be difficult to incorporate commercial actors into official multilateral settings. Nonetheless, the international community should strive to bring commercial actors into conversations about responsible innovation and normative frameworks that proactively mitigate the risks of emerging technologies.

Conclusion

While the future of emerging technologies is uncertain, it is clear that a wealth of technological developments will impact the future of geopolitical cooperation and competition, avenues of crisis stability and conflict escalation control, and arms control verification and compliance tools. Policymakers today are beginning to wrestle with these difficult questions. However, governing technologies that have not yet fully come to fruition poses logistical challenges. Many countries may be hesitant to engage in meaningful arms control to limit the deployment of these new tools prior to knowing the full implications of a technology or set of technologies.

As policymakers engage these challenges, it is crucial to consider who will be involved in the conversation and where these conversations should occur to maximize impact. A multi-stakeholder approach that includes industry and true interdisciplinary expertise will be required to gain a comprehensive understanding of the ways that additive technological developments may overlap and impact each other. These conversations should take place in multiple international fora, including within the NPT. Many of these technologies have direct impacts for the future of nuclear strategy and non-proliferation, so the NPT must grow and address these new challenges to remain effective.

Jessica Link and Joseph Rodgers
Wilton Park | March 2023