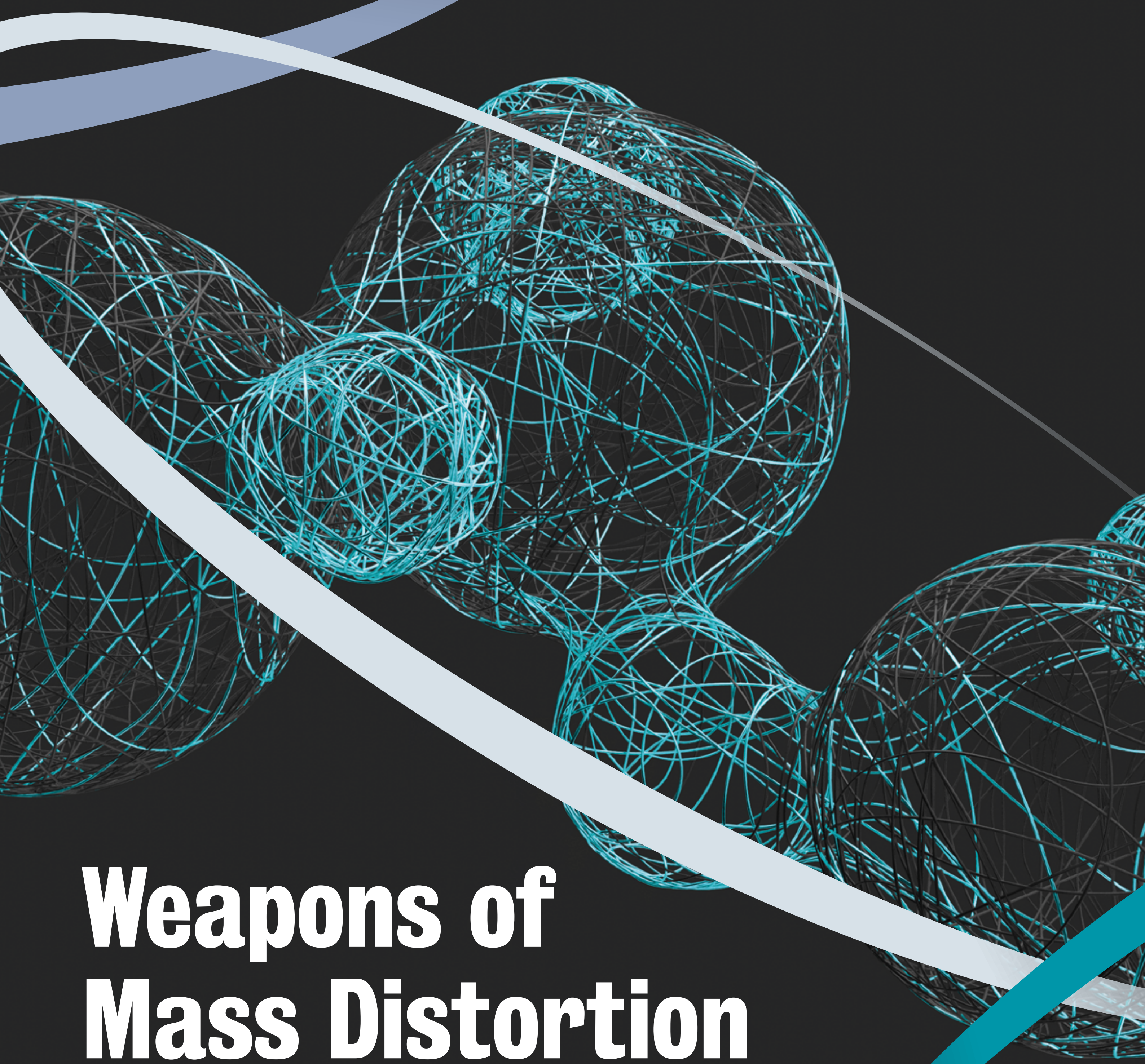
CENTRE FOR SCIENCE
& SECURITY STUDIES

KING'S
College
LONDON



Weapons of Mass Distortion

A new approach to emerging technologies, risk reduction, and the global nuclear order

By Marina Favaro

2021 EDITION

Table of contents

Biography: Marina Favaro

Marina Favaro is a Consultant at the Centre for Science and Security Studies (CSSS) at King's College London and a Research Fellow at the Institute for Peace Research and Security Policy (IFSH) at the University of Hamburg. Her research focuses on the impact of emerging technologies on arms control. From 2020 to 2021, Marina managed the Emerging Technologies research programme at the think tank BASIC. Before that, Marina worked as an Analyst at RAND Europe, where her research focused on space security, cybersecurity, defence innovation, and the impact of emerging technologies on society. Marina conducts quantitative and qualitative research through a variety of methods, including futures and foresight methods (eg horizon scanning, STREAM, Delphi, and scenario development). She holds a Master's degree in international relations and politics from the University of Cambridge.

About the Centre for Science and Security Studies at King's College London

The Centre for Science and Security Studies (CSSS) is a multi-disciplinary research and teaching group at King's College London that brings together scientific experts with specialists in politics, international relations, and history. CSSS forms part of the School of Security Studies at King's and draws on experts from the Department of War Studies and the Department of Defence Studies. Members of the Centre conduct scholarly and policy-relevant research on weapons proliferation, non-proliferation, verification and disarmament, nuclear security, space security and mass effect terrorism including the CBRN (chemical, biological, radiological and nuclear) dimension. In addition to academic staff, CSSS hosts masters and postgraduate research students, as well as visiting fellows and associates drawn from the academic, government, and business sectors. Our educational activities include contributions to the undergraduate and postgraduate offerings in the Department of War Studies, as well as professional development workshops for industry professionals.

Executive summary

Emerging technologies are changing the nuclear landscape in three ways. First, technological change is accelerating and originates in the private sector. In recent years, the locus of innovation has shifted away from militaries and governments and towards the private sector, which could decrease the public sector's awareness of, or control over, how new technologies mature and are applied. Second, nuclear policy is struggling to keep pace with emerging technologies. Although many policymakers and experts agree that emerging technologies increase the risks of nuclear use, multilateral institutions have been slow to address the subject of emerging technologies, because of the complexity of the issue and already-crowded agendas. Third, nuclear risks are rising, yet there is no clear path forward for how nuclear possessors and non-possessors can cooperate to make progress on nuclear risk reduction.

Notwithstanding the heterogeneity and lack of definitional clarity of 'emerging technologies', policymakers tend to treat them as a broad risk category. This study helps policymakers to better allocate a state's limited resources by asking: Which emerging technologies are most likely to escalate a crisis? How can policymakers and scholars better understand the ways in which nuclear weapon states might feel this impact? And which nuclear risk reduction measures could mitigate any potential risks of emerging technologies, while also capitalising on the benefits and opportunities that they present?

To answer these questions, this study uses a mixed-methods approach that combines qualitative and quantitative techniques. Using data generated by a technology scoring exercise, this study groups 10 shortlisted technologies into four technology clusters, using Machine Learning (ML). The four technology clusters comprise technologies that 1) distort, 2) compress, 3) thwart, and 4) illuminate in a crisis. Cluster 1 encompasses technologies that are capable of interrupting data flows and 'distorting' the information landscape. Technologies in Cluster 2 impact the speed of conflict and could 'compress' decision-making timelines. The defining feature of Cluster 3 is its ability to credibly 'thwart' or blunt a nuclear attack. Finally, Cluster 4 'illuminates' insofar as it provides more accurate and comprehensive data flows to decision-makers. For each technology cluster, this paper identifies fit-for-purpose risk reduction measures.

> **This study groups 10 shortlisted technologies into four technology clusters, using Machine Learning. The four technology clusters comprise technologies that distort, compress, thwart, and illuminate in a crisis**

Of the four technology clusters, experts determined that technologies in Cluster 1 are the most concerning in terms of nuclear risk, due to their potentially high impact and the high feasibility of their implementation. To combat the nuclear risks created by Cluster 1, this paper suggests that the United Kingdom and other nuclear possessors should use non-consensual deep fake pornography as a legislative entry point for prohibiting certain uses of a digital replica of a person. Furthermore, this paper suggests that nuclear weapon states should champion the protection of space-based assets linked to early warning as part of the UN General Assembly's Draft Resolution on responsible space behaviours. This paper also proposes broader risk reduction recommendations, with relevance to states with and without nuclear weapons. Foremost amongst these is the recommendation that the public sector in nuclear weapon states must cooperate more closely with the private sector around the potential harms of dual-use technologies. Otherwise, commercial off-the-shelf technologies could become an unwitting part of another state or non-state actors' foreign policy objectives. Finally, this paper recommends that nuclear possessors, such as the UK, should expand the nature of existing partnerships with non-possessors, to include, for example, Estonia and the Netherlands, who have specialist knowledge in cybersecurity.

Bringing emerging technologies into the ongoing discussion about nuclear risk reduction is necessary to ensure that existing institutions, such as the Nuclear Non-Proliferation Treaty (NPT) remain relevant. More importantly, addressing emerging technologies in multilateral institutions like the NPT is necessary to reduce the risks of a catastrophic nuclear exchange.

The blind spot of the global nuclear order

In June 2020, United Nations (UN) High Representative for Disarmament, Izumi Nakamitsu, dedicated a keynote address to emerging technologies and nuclear risk. She stated, 'developments in a variety of technologies are diminishing predictability, shared understandings and trust, while raising the risks of misperception, arms races, and potential escalation through miscalculation.' In particular, she stressed, 'none of the nuclear weapons-related forums are discussing the intersection between technology and nuclear risk, adding to decreasing transparency and a climate of misperception.'[1] On the one hand, many policymakers and experts agree that emerging technologies increase the risks of nuclear use.[2] On the other hand, multilateral institutions have been slow to address the subject of emerging technologies, because of the complexity of the issue and already-crowded agendas.

The impact of emerging technologies on nuclear weapons risks, as well as prospects for escalation, arms control, and disarmament, are of increasing concern for states in the global nuclear order. Nuclear policymakers must address the rapid pace of technological change; but first, they need to know which emerging technologies are likely to impact nuclear risk, in what ways, and what can be done to mitigate these risks while remaining cognisant of the potential benefits of innovation.

Past efforts to address the intersection of emerging technologies and nuclear risk through existing institutions have been patchy. Part of the reason for this is definitional. The term 'emerging technologies' is excessively broad, making it difficult for policymakers to decide which risks to focus on. Policymakers need a new way of thinking and talking about emerging technologies and nuclear risk. This study creates common parameters for measuring the future impact of emerging technologies on crisis stability. It thereby helps to 'future-proof' existing institutions and mechanisms, such as the Nuclear Non-Proliferation Treaty (NPT).[3]

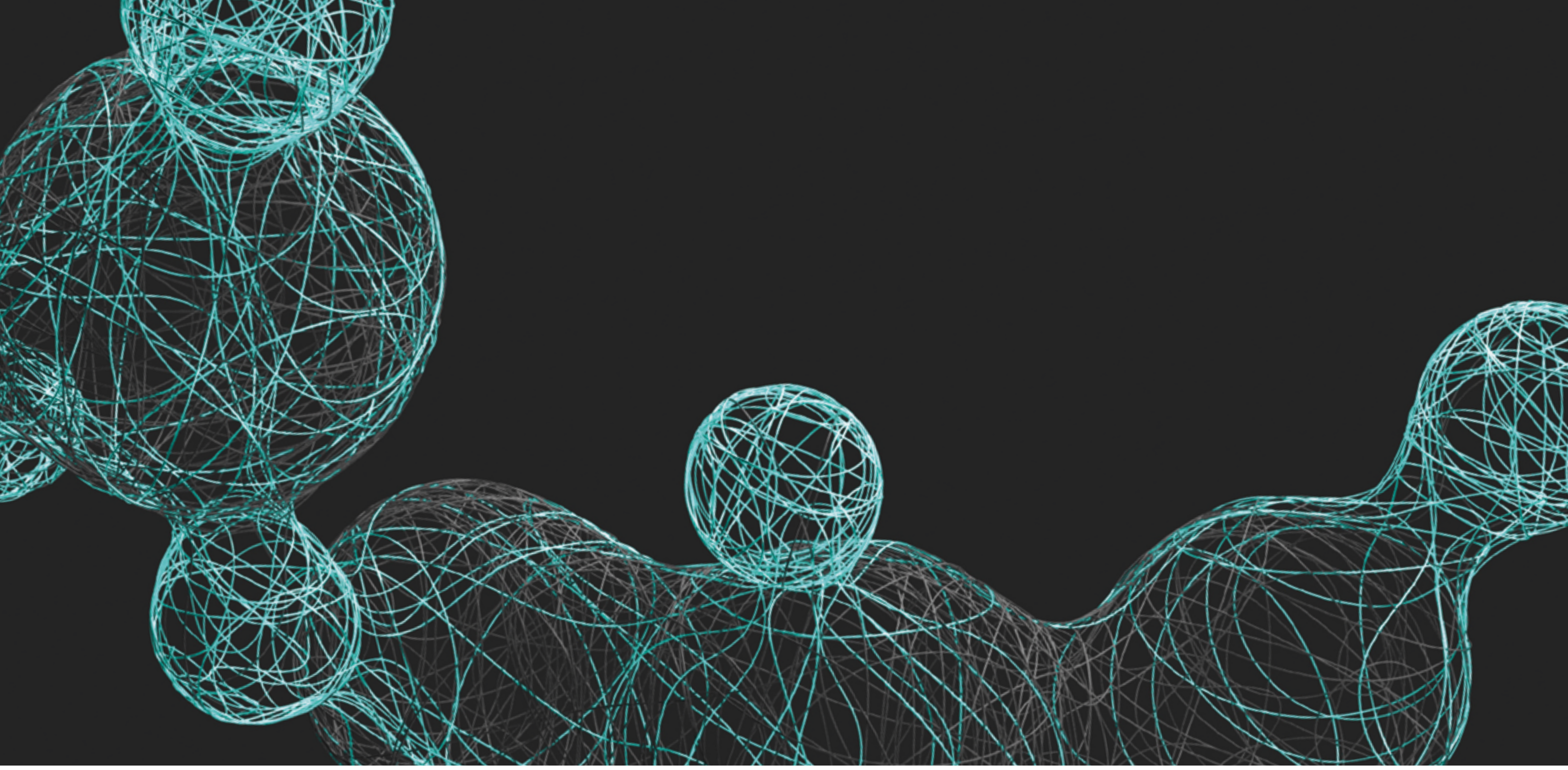
This study adapts the North Atlantic Treaty Organisation (NATO)'s 2020 definition for emerging technologies to denote 'those technologies or scientific discoveries that are expected to reach maturity in the period 2020-[2030]; and are not yet widely in use or whose effects on defence [and] security are not entirely clear.'[4] But which emerging technologies are most likely to impact crisis stability in the next 10 years? How can policymakers and scholars alike better

[1] Izumi Nakamitsu, Keynote Speech at the Virtual UK Project on Nuclear Issues 2020 Annual Conference Royal United Services Institute for Defence and Security Studies, June 10, 2020, available at: https://front.un-arm.org/wp-content/uploads/2020/06/10-June-High-Representative-Keynote-at-RUSI-UK-PONI-Annual-Conference-2020.pdf

[2] Caitlin Talmadge, 'Emerging technology and intra-war escalation risks: Evidence from the Cold War, implications for today', *Journal of Strategic Studies,* 42:6 (2019), available at: https://www.tandfonline.com/doi/abs/10.1080/01402390.2019.1631811

[3] Future-proofing is the process of anticipating the future and developing methods of minimising the effects of shocks and stresses of future events.

[4] NATO Science & Technology Organization, 'Science & Technology Trends 2020-2040: Exploring the S&T Edge', March 2020, available at: https://www.nato.int/nato_static_fl2014/assets/pdf/2020/4/pdf/190422-ST_Tech_Trends_Report_2020-2040.pdf



understand the ways in which nuclear weapon states might feel this impact? And which nuclear risk reduction measures could mitigate the potential risks of emerging technologies whilst allowing states to capitalise on the benefits and opportunities presented by innovation?

To answer these questions, this paper applies a novel research method, the Systematic Technology Reconnaissance, Evaluation and Adoption Methodology (STREAM),[5] to identify which emerging technologies might be most destabilising in a crisis in the next 10 years, using the example of the United Kingdom (UK). In other words, this study asks whether a given technology is likely to increase, or reduce, the potential of a crisis escalating past the nuclear threshold in a theoretical crisis involving the UK and another nuclear possessor. The findings, however, apply to a wider range of actors, not just nuclear possessors.

> “The most concerning technologies in terms of nuclear risk may be those that ‘distort’ the information space”

The study has three primary findings. First, rather than looking at ‘emerging technologies’ as a broad risk category, the study proposes an original framework with four technology clusters; those that 1) distort, 2) compress, 3) thwart, and 4) illuminate in a crisis. Second, based on the study findings, the most concerning technologies in terms of nuclear risk may be those that ‘distort’ the information space, including satellite spoofing or deep fake technologies, with adverse effects for trust and online civic culture. Experts involved in the STREAM analysis assessed the technologies in this cluster as both of high impact and highly likely to be developed by the UK and other international actors, over the next ten years. Third, the findings suggest that a priority for nuclear possessors and non-possessors alike in reducing nuclear risk will be to find new ways of engaging with the private sector, which is ultimately responsible for developing many of the technologies discussed herein.

This paper begins by discussing the need for a new approach to the intersection of nuclear risk and emerging technologies and outlines the research design. It then discusses the four technology clusters with fit-for-purpose nuclear risk reduction measures. Finally, it offers recommendations for a variety of actors to address risks associated with emerging technologies across the four clusters. One of the most important recommendations is that nuclear weapon states (NWS) and non-nuclear weapon states (NNWS) should cooperate to identify ways to utilise these technologies to reduce nuclear risk, rather than seeing them all as essentially destabilising.

[5] STREAM is a RAND-developed method was originally developed by Popper et al. (2013) for application in the transportation sector. It has since been applied in a range of security and defence topics, but STREAM has never, to the author’s knowledge, been used in the nuclear policy space.

[6] Declan Butler, ‘Tomorrow’s World’, Nature, 2016, available at: https://www.nature.com/news/polopoly_fs/1.19431!/menu/main/topColumns/topLeftColumn/pdf/530398a.pdf?origin=ppub

Towards a new approach to emerging technologies

The strategic landscape has changed significantly since 2015 and the last NPT Review Conference (RevCon), but three trends in particular may shape the future of the global nuclear order and warrant a closer look: the pace of technological change, the impact of emerging technologies on nuclear policy, and rising nuclear risks.

First, technological change is rapid and originates in the private sector. Innovation has long been a focal point in military competition, but some futurists contend that the pace of technological change is accelerating,[6] with traditional defence acquisition and international legal structures struggling to keep up. The UK government's recent review of security and defence policy, 'Global Britain in a Competitive Age: The Integrated Review of Security, Defence, Development and Foreign Policy' (2021) asserts that, 'The [science & technology] landscape has changed significantly since 2015 and the pace of change will accelerate further to 2030.'[7] The Integrated Review mentions, in particular, artificial intelligence (AI), quantum and data analytics.[8] There has also been a shift in the locus of innovation away from militaries and governments and towards the private sector. The Defence and Security Accelerator (DASA) in the UK Ministry of Defence affirms that, 'technological innovation is now more likely to come from the private sector, from companies based in other countries, outside of government's control, and where the interests of one government or another are of very minority interest.'[9] This has resulted in widespread applications of dual-use technologies with both military and civil applications. A prominent example of this is AI, which is used for facial recognition in retail stores,[10] as well as for processing intelligence, surveillance, target acquisition and reconnaissance information for national security and defence purposes.[11] Collectively, such trends raise questions about how governments can best work with private sector organisations.

Second, nuclear policy is struggling to keep pace with emerging technologies. Emerging technologies will change the force posture of states with and without nuclear weapons. Consider, for example, how aggressive technology programmes could enhance the ability of NNWS to impact strategic stability, by diminishing predictability and increasing the risks of escalation through miscalculation. Addressing the risks of emerging technologies in the nuclear realm will require a multilateral effort by nuclear possessors and non-possessors, through the NPT. Given that a perceived lack of progress towards nuclear disarmament is polarising international perceptions of the NPT, cooperation on emerging technologies and nuclear risk reduction offers an important and timely opportunity for NWS and NNWS. Cooperation between these groupings would help to achieve the Treaty's ultimate goals of disarmament, non-proliferation, and peaceful uses of nuclear energy, while reducing the likelihood and impact of a catastrophic nuclear exchange. Indeed, to continue to ignore emerging technologies risks rendering the NPT irrelevant.[12] More importantly, to continue to ignore emerging technologies could increase the risk of nuclear weapons use.

Third, nuclear risks are rising. A 2019 report by the UK's House of Lords concludes that 'The risk of the use of nuclear weapons has increased, in the context of rising inter-state competition, a more multipolar world, and the development of new capabilities and technologies.'[13] Given this challenge, states are

[7] HM Government, 'Global Britain in a competitive age: The Integrated Review of Security, Defence, Development and Foreign Policy', March 2021, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age-_the_Integrated_Review_of_Security__Defence__Development_and_Foreign_Policy.pdf, p. 30.

[8] Ibid.

[9] Defence and Security Accelerator, 'Future technology trends in security', available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/728113/Future_trends_research_V6.pdf

[10] Matt Burgess, 'Some UK Stores Are Using Facial Recognition to Track Shoppers', WIRED, December 20, 2020, available at: https://www.wired.com/story/uk-stores-facial-recognition-track-shoppers/

[11] David Vergun, 'Artificial Intelligence Key to Maintaining Military, Economic Advantages, Leaders Say', U.S. Department of Defence, April 9, 2021, available at: https://www.defense.gov/Explore/News/Article/Article/2567486/artificial-intelligence-key-to-maintaining-military-economic-advantages-leaders/

[12] Heather Williams, 'Remaining relevant: Why the NPT must address emerging technologies', King's College London, August 2020, available at: https://www.kcl.ac.uk/csss/assets/remaining-relevant-new-technologies.pdf

increasingly focused on nuclear risk reduction through various collaborative efforts; however, as is the case with emerging technologies, there is no shared definition for nuclear risk. As a result, different actors have different conceptualisations of nuclear risk reduction. For NWS, it means reducing the risk of nuclear use, whether accidental or deliberate. For NNWS, nuclear disarmament and the elimination of nuclear weapons is the best way to reduce nuclear risks. Between these two extremes, there are a range of policy options, which include de-alerting, de-targeting, taking weapons out of operational service, 'no first use' commitments, further reductions of stockpiles, more transparency about postures and actual use scenarios, and confidence-building measures.[14] One objective of the Stockholm Initiative on Nuclear Disarmament[15] and the NPT, for example, is to identify 'ways of enhancing transparency and of reducing risks of any use of nuclear weapons.'[16] Another initiative, Creating an Environment for Nuclear Disarmament (CEND),[17] includes a subgroup specifically devoted to risk reduction, co-chaired by Finland and Germany.[18] But the theory and practice of risk reduction remain relatively broad and thus potentially difficult to translate into concrete action.

While there is partial and temporary consensus on the importance of risk reduction efforts, there is no clear path forward for how NWS and NNWS can cooperate to make progress on this issue. Wilfred Wan of the United Nations Institute for Disarmament Research (UNIDIR) has offered one particularly helpful approach, looking at five potential scenarios of nuclear use: doctrinal, escalatory, unauthorised, accidental, and interactive.[19] But in the context of emerging technologies, which risks are the most likely to materialise in the next 10 years? And which could have the biggest impact, including across these different scenarios?

This study utilises the STREAM method to evaluate dissimilar technologies. STREAM culminates in a technology scoring exercise, which tasks an interdisciplinary mix of subject-matter experts with assessing the impact that a given technology might have on a given function, as well as any barriers to its implementation in a specific context, over a defined timeframe.[20] Three components of the research design are important to note at the outset: the survey that constitutes the core of the STREAM process was specifically designed to evaluate technologies from the perspective of UK decision-makers; the analysis was primarily interested in the impact of technologies on crisis escalation; and the survey respondents evaluated a shortlist of 10 technologies, derived from a much longer list of candidate technologies. Each of these factors is briefly explained below, but *Annex A: Data collection, analysis, and validation* contains a more complete methodology.

Focusing on the United Kingdom

This study uses the United Kingdom to ground its discussion about the impact of emerging technologies on nuclear risk. The UK is a relevant focus of discussion because it is a vocal proponent of nuclear risk reduction, is a NWS with a proven track record in creating partnerships with NNWS, and has ambitions to become a global 'science & technology (S&T) superpower'.[21] Helpfully, the government's recent Integrated Review also emphasises two significant ways in which emerging technologies are impacting the UK's nuclear deterrent.

First, the Integrated Review states that, 'in recognition of the evolving security environment, including the

[13] 'Rising nuclear risk, disarmament, and the Nuclear Non-Proliferation Treaty', UK House of Lords Select Committee on International Relations, April 24, 2019, available at: https://publications.parliament.uk/pa/ld201719/ldselect/ldintrel/338/33803.htm

[14] Alexander Kmentt, 'Nuclear deterrence perpetuates nuclear risks: the risk reduction perspective of TPNW supporters', European Leadership Network, December 4, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/nuclear-deterrence-perpetuates-nuclear-risks-the-risk-reduction-perspective-of-tpnw-supporters/

[15] The Stockholm Initiative was launched in 2019 with the aim of strengthening disarmament diplomacy within the context of the Non-Proliferation Treaty.

[16] Government Offices of Sweden, 'The Stockholm Ministerial Meeting on Nuclear Disarmament and the Non-Proliferation Treaty', June 11, 2019, available at: https://www.government.se/statements/2019/06/the-stockholm-ministerial-meeting-on-nuclear-disarmament-and-the-non-proliferation-treaty/

[17] The Creating an Environment for Nuclear Disarmament (CEND) initiative is an informal Track 1 dialogue with dozens of state participants working on three main topics: 1) reducing reliance on nuclear weapons; 2) mechanisms and institutions; and 3) risk reduction.

[18] Heather Williams, 'CEND and a changing global nuclear order', European Leadership Network, February 18, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/cend-and-a-changing-global-nuclear-order/

[19] Wilfred Wan, 'Nuclear Risk Reduction: A framework for analysis', UNIDIR, 2019, available at: https://unidir.org/publication/nuclear-risk-reduction-framework-analysis

[20] Steven W. Popper, Nidhi Kalra, Richard Silberglitt, Edmundo Molina-Perez, Youngbok Ryu, and Michael Scarpati, 'Strategic Issues Facing Transportation, Volume 3: Expediting Future Technologies for Enhancing Transportation System Performance', NCHRP Report 750, 2013, available at: https://doi.org/10.17226/22448

[21] HM Government, 'Global Britain in a competitive age: The Integrated Review of Security, Defence, Development and Foreign Policy', March 2021, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age-_the_Integrated_Review_of_Security__Defence__Development_and_Foreign_Policy.pdf

[22] Ibid.

[23] United Nations Office of Disarmament Affairs, 'Treaty on the Non-Proliferation of Nuclear Weapons (NPT)', available at: https://www.un.org/disarmament/wmd/nuclear/npt/

developing range of technological and doctrinal threats' the UK will no longer commit to reaching a ceiling of 180 nuclear warheads by the mid-2020s, and instead will set a cap of 260 operational warheads in the stockpile.[22] The invocation of technological threats as a primary justification for the UK's reassessment illustrates how emerging technologies are directly impacting the global nuclear order and the NPT in particular, given that the UK is among the 191 States parties to the NPT who have committed to the 'cessation of the nuclear arms race' and 'general and complete disarmament'.[23] This decision will have ramifications for the UK's credibility as a leader in transparency and disarmament at the 2021 NPT Review Conference.[24]

Second, the Integrated Review states that the UK 'reserve[s] the right to review its assurance [not to use, or threaten to use, nuclear weapons against any non-nuclear weapon state party to the NPT] if the future threat of weapons of mass destruction, such as chemical or biological capabilities, or emerging technologies that could have a comparable impact, makes it necessary.'[25] The government later clarified that the UK 'are not considering using [their] nuclear deterrent to deter cyber-attacks.'[26] Nevertheless, Tom Plant, an expert at the Royal United Services Institute, points out that the change of language in the Integrated Review is significant.[27] It suggests that, in the future, new technologies and their associated behaviours could create unprecedented nuclear risks. Furthermore, it acknowledges that emerging technologies, particularly when used in combination with each other, could rival weapons of mass destruction in their strategic or tactical impact. Above all, the Integrated Review recognises that 'rapid technological change' will be 'of central significance to the strategic context'.[28]

> **This study utilises the STREAM method to evaluate dissimilar technologies**

Beyond the headline-grabbing pronouncements, some scholars argue that the Integrated Review lacks depth and detail, especially in relation to emerging technologies.[29] This illustrates a tendency of policymakers in the UK and globally to treat emerging technologies as a broad risk category, notwithstanding their heterogeneity. So-called emerging technologies span multiple operating domains,[30] are at different maturity or Technology Readiness Levels (TRL),[31] have different barriers to implementation, and will impact different elements of the global nuclear order to varying extents and in varying timescales. Policymakers need a way to compare different technologies in terms of common parameters to determine where a state should allocate its limited resources.

Focusing on crisis stability

This study focuses on the ability of technologies to impact crisis stability. A crisis is defined as 'stable' if neither side has or perceives an incentive to use nuclear weapons first out of the fear that the other side is about to do so.[32] This common description reflects a bipolar, Cold War dynamic, with scholars still divided over how best to update and extend this model to reflect the

[24] Heather Williams, 'U.K. Nuclear Weapons: Beyond the numbers', War on the Rocks, April 6, 2021, available at: https://warontherocks.com/2021/04/u-k-nuclear-weapons-beyond-the-numbers/

[25] HM Government, 'Global Britain in a competitive age: The Integrated Review of Security, Defence, Development and Foreign Policy', March 2021, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age-_the_Integrated_Review_of_Security__Defence__Development_and_Foreign_Policy.pdf

[26] HM Government, 'The 2021 Integrated Review: nuclear frequently asked questions', 27 April 2021, available at: https://www.gov.uk/guidance/the-2021-integrated-review-nuclear-frequently-asked-questions

[27] CNBC, 'Britain changes policy so it can use nuclear weapons in response to 'emerging technologies', March 17, 2021, available at: https://www.cnbc.com/2021/03/17/uk-changes-policy-so-it-can-use-nukes-in-response-to-emerging-tech.html

[28] HM Government, 'Global Britain in a competitive age: The Integrated Review of Security, Defence, Development and Foreign Policy', March 2021, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age-_the_Integrated_Review_of_Security__Defence__Development_and_Foreign_Policy.pdf

[29] Alexi Drew, 'The Integrated Review: a strategy of 'bare bones' yet to be fleshed out', King's College London, March 20, 2021, available at: https://www.kcl.ac.uk/news/the-integrated-review-a-strategy-of-bare-bones-yet-to-be-fleshed-out

[30] The UK currently recognises five operating domains in its defence doctrine ('land', 'air', 'maritime', 'space', and 'cyber and electromagnetic') but NATO and other individual nations use varying definitions.

[31] TRL is a concept originally developed by the National Aeronautics and Space Administration (NASA) that provides a useful shorthand for interpreting technology maturity level. The 1-9 scale can be partitioned into TRL 1-3: Research, TRL 4-6: Development, and TRL 7-9: Deployment. Source: NASA, 'Technology Readiness Level', October 28, 2012, available at: https://www.nasa.gov/directorates/heo/scan/engineering/technology/technology_readiness_level

[32] James M. Acton, 'Reclaiming Strategic Stability', Carnegie Endowment for International Peace, February 5, 2013, available at: https://carnegieendowment.org/2013/02/05/reclaiming-strategic-stability-pub-51032

added complexities of today's strategic environment. To modernise this Cold War conceptualisation, James Acton suggests that 'first strike stability is… a necessary, but not sufficient, condition for crisis stability'.[33] In other words, whereas first strike stability focuses solely on the technical characteristics of each side's strategic forces (eg the hardening of silos, the accuracy of missiles, the effect of missile interceptors, etc), crisis stability ought to be defined more broadly.

This study understands crisis stability to refer to a scenario in which 'emotion, uncertainty, miscalculation, misperception, or the posture of forces' do not incentivise leaders 'to strike first, to avoid the worse consequences of incurring a first strike'.[34] This definition emphasises that in addition to technical characteristics, other psychological, political and strategic factors are also relevant to crisis stability.[35] These factors are more difficult to quantify and model than the technical characteristics of weapon systems. Scholars Kristin Ven Bruusgaard and Jaclyn A. Kerr highlight an additional complication to the contemporary definition for crisis stability by adding that 'the current information environment presents additional challenges for retaining stability in crisis'.[36] This includes new tools of dis- and misinformation and an abundance of unverified data that is available to decision-makers.

> **STREAM helps to identify how the use of a given technology will impact the stability of a crisis between the UK and another nuclear actor, and what is the projected trajectory for the development of this technology by the UK**

As a final point on crisis stability, Aaron R. Miles clarifies that, 'Crisis stability does not imply that crises are impossible or will be infrequent. Crisis stability means that when crises do arise, the system does not drive them to get worse.'[37] Put differently, the central concern of crisis stability is not necessarily to deter conflict, but rather to avert the escalation of an ongoing crisis past the nuclear threshold.[38] Thus, this study does not ask which technologies would be most likely to catalyse a crisis in the first instance, but rather, which technologies might have the potential to exacerbate a crisis once it has already begun.

Using a mixed-methods approach

This study shortlisted 10 technologies for consideration, based on a literature review, key informant interviews, a proof-of-concept study, and a Red Team workshop:

- AI-powered cyber operations
- AI for intelligence, surveillance, and reconnaissance (ISR)
- Deep-fake technology
- Directed energy weapons
- Hypersonic missiles
- Kinetic anti-satellite (ASAT) capabilities
- Rendezvous and Proximity Operations (RPO) in space
- Satellite jamming and spoofing systems
- Small satellites ('smallsats') for ISR
- Swarm robotics

The study used a mixed-methods approach that combines qualitative and quantitative techniques for data collection and analysis (see *Annex A* for full detail on the research methods).[39] As introduced above, the STREAM method evaluates the potential impact of each technology on a given function and what barriers to implementation exist in a specific context. Put differently, STREAM helps to identify how the use of a given technology will impact the stability of a crisis between the UK and another nuclear actor, and what is the projected trajectory for the development of this technology by the UK.

[33] Ibid.

[34] Glenn A. Kent, David E. Thaler, 'First-Strike Stability: A Methodology for Evaluating Strategic Forces', RAND Project Air Force, August 1989, available at: https://www.rand.org/pubs/reports/R3765.html

[35] Ibid.

[36] Kristin Ven Bruusgaard and Jaclyn A. Kerr, 'Crisis Stability and the Impact of the Information Ecosystem', Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict, March 15, 2020, available at: https://www.hoover.org/sites/default/files/research/docs/trinkunas_threetweetstomidnight_137-158_ch.7.pdf (emphasis in original)

[37] Aaron R. Miles, 'The dynamics of strategic stability and instability', *Comparative Strategy,* 35:5 (2016), available at: https://www.tandfonline.com/doi/abs/10.1080/01495933.2016.1241005, p. 425.

[38] Kristin Ven Bruusgaard and Jaclyn A. Kerr, 'Crisis Stability and the Impact of the Information Ecosystem', Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict, March 15, 2020, available at: https://www.hoover.org/sites/default/files/research/docs/trinkunas_threetweetstomidnight_137-158_ch.7.pdf (emphasis in original)

[39] This study exhibits a degree of 'methods pioneering' by combining methods that, to the author's knowledge, have never been applied in the nuclear policy community.

An interdisciplinary group of experts answered fourteen questions about each of the 10 shortlisted technologies. In the technology scoring exercise, respondents assigned numerical values to different aspects of crisis stability (eg ability to deliver a disarming first strike, ability to increase mis/disinformation), as well as barriers to implementation (eg budgetary, human, regulatory, ethical, legal, technical) for the UK, as a proxy for other states.[40] They also provided qualitative comments in the technology scoring exercise and via email. In total, 61 subject-matter experts responded to the STREAM technology scoring exercise. Machine Learning was used to analyse the resulting data. This revealed four clusters of data.[41]

There are a few important caveats to this study that warrant a brief exploration.

First, STREAM considers technologies individually, when in fact, disruptive effects will most likely occur through combinations of technologies (both existing and emerging) and the complex interactions between them.[42]

Second, it is very difficult to capture game theoretic dynamics in a technology scoring exercise. STREAM assesses how impactful a given technology might be, and the likelihood of its use, by a given actor (here, the UK). Evaluating the feasibility of implementing that technology for 'red' (meaning adversarial) forces falls outside the scope of this study. In other words, this study does not account for the ways in which implementation drivers and barriers might be similar or different between the UK and its adversaries, nor does it account for future countermeasures.

Third, this method risks reifying the 'hype' surrounding certain technologies, because even subject-matter experts are susceptible to various cognitive biases when projecting the uncertain future of how a new technology might develop or find real-world application.[43] This could manifest in two ways: experts could score a given technology as higher impact than it really is and/or more experts could score a given technology, at the expense of other technologies.

Fourth, this study characterises certain technologies as 'lower feasibility' relative to the rest of the shortlist. This should be understood in relative, not absolute terms. Given the 10-year time horizon, many technologies on the shortlist are *already* operational in the UK or elsewhere, to some degree.[44]

[40] The respondents (n=61) display nearly equal representation from the policy community and the S&T community (58 per cent:42 per cent), and the numerical values in the survey correspond to qualitative descriptors, as on a Likert scale.

[41] The use of Machine Learning to cluster the technologies sets this study apart from that which came before it. Whereas prioritisation uses the mean of all impact and implementation scores (wherein each criterion is weighted equally), clustering groups technologies that scored similarly across various criteria. To work off the mean of the impact and implementation scores would generate a priority list of most-to-least impactful technologies but would lose the detail of *in what ways* the technologies impact crisis stability. Clustering is a less prescriptive approach than prioritisation because it allows policymakers to decide the relative weight that they will ascribe to each criterion (ie are they more concerned with technologies that will impact crisis stability variable X or crisis stability variable Y?). Clustering the technologies generates meaningful policy outputs by preserving more nuance in the expert scores.

[42] NATO Science & Technology Organization, 'Science & Technology Trends 2020-2040: Exploring the S&T Edge', March 2020, available at: https://www.nato.int/nato_static_fl2014/assets/pdf/2020/4/pdf/190422-ST_Tech_Trends_Report_2020-2040.pdf

[43] Stephen M. Meyer's *The Dynamics of Nuclear Proliferation* (1984) is a powerful illustration of what subject-matter experts can demonstrate. Meyer successfully debunked the technological imperative thesis by providing the first empirical, statistical model explaining why particular countries became nuclear powers when they did. What is remarkable about this study, apart from being the first of its kind, is that experts answered specific questions in a way that showed they did not believe that technology drives proliferation. This contradicted the technological imperative thesis, which was accepted by large segments of the academic and policy communities at the time.

[44] The question remains: To what degree? For example, 'narrow' AI is currently used in a military context and air defence systems have exhibited some level of both automation and autonomy since the 1970s. Artificial Narrow Intelligence, also sometimes known as 'weak' AI, has been programmed to perform a single task, such as checking the weather, playing chess, natural language processing, recognising/classifying objects, or augmenting search engines. Artificial General Intelligence (AGI) or 'strong' AI is still a way away, with some scholars disputing whether it is even feasible. AGI refers to machines that exhibit human intelligence, enabling them to successfully perform any intellectual task (eg reasoning, problem solving, making decisions under uncertainty, innovation, etc) that a human can. So, the use of narrow AI for cyber operations is high feasibility/high TRL, whereas the use of 'general' AI for the same application is low feasibility/low TRL. This illustrates the difficulty of evaluating a technology that is AI-enabled because this relies on a forecasting of what AI might be capable of in 10 years' time. Source: Ingvild Bode and Tom Watts, 'Meaning-less Human Control: Lessons from air defence systems on meaningful human control for the debate on AWS', University of Southern Denmark, February 2021, available at: https://dronewars.net/wp-content/uploads/2021/02/DW-Control-WEB.pdf; Jamie Berryhill, Kévin Kok Heang, Rob Clogher, and Keegan McBride, 'Hello, World: Artificial intelligence and its use in the public sector', OECD, November 2019, available at: https://oecd-opsi.org/wp-content/uploads/2019/11/AI-Report-Online.pdf; Congressional Research Service, 'Artificial Intelligence and National Security', November 10, 2020, available at: https://fas.org/sgp/crs/natsec/R45178.pdf

Technology clusters: Distort, compress, thwart, and illuminate

Using data generated by the STREAM technology scoring exercise, this study groups the 10 technologies into four technology clusters, based on the challenges and opportunities that they pose to crisis stability. The four technology clusters are: 1) distort, 2) compress, 3) thwart and 4) illuminate. The technology clusters are visualised in *Figure 1: Technologies' impacts on crisis stability* and summarised in *Table 1: The four technology clusters.* In Figure 1, feasibility of implementation is on the x-axis, impact is on the y-axis and the bubble size corresponds with TRL, meaning that the larger the bubble, the more mature the technology is presently.

Figure 1: Technologies' impacts on crisis stability

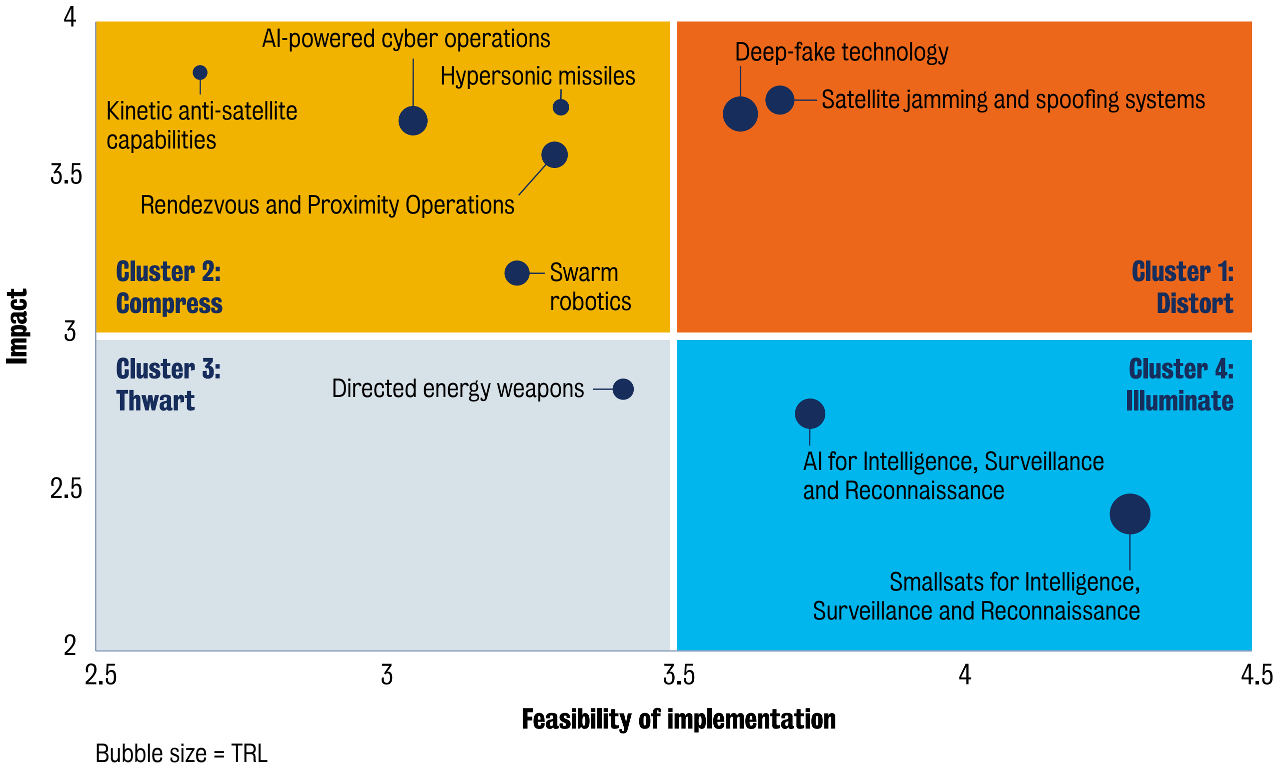


Table 1: The four technology clusters

	Name	Impact	Feasibility of implementation	Technologies
Cluster 1	"Distort"	Higher	Higher	Satellite jamming and spoofing
				Deep fake technology
Cluster 2	"Compress"	Higher	Lower	Hypersonic missiles
				Swarm robotics
				Kinetic anti-satellite capabilities
				AI-powered cyber operations
				Rendezvous and Proximity Operations
Cluster 3	"Thwart"	Lower	Lower	Directed energy weapons
Cluster 4	"Illuminate"	Lower	Higher	AI for Intelligence, Surveillance and Reconnaissance
				Smallsats for Intelligence, Surveillance and Reconnaissance

These four technology clusters warrant closer examination in turn, along with potential risk reduction measures unique to their applications.

Cluster 1 – Distort

Experts assessed these technologies as higher impact and higher feasibility of implementation. Deep fake technology, satellite jamming, and satellite spoofing systems are capable of interrupting data flows and disrupting the information landscape which may, in turn, lead to increased misperception, confusion, and uncertainty in a crisis. In addition to increasing dis- and misinformation during a crisis, experts assessed that these technologies could also reduce decision-making time and situational awareness during a crisis and could erode Nuclear Command, Control, and Communications (NC3). Deep fakes use Machine Learning (ML) techniques to manipulate audio and video, with the objective of creating realistic forgeries.

Recent advances in ML could enhance the effectiveness of malicious media manipulation efforts.[45] The utility of deep fake techniques is not confined to one geographic region or a single adversary and could become an attractive tactic to gain an asymmetric advantage by state and non-state actors alike. Even if deep fakes do not convince citizens, they may sow uncertainty which could, in turn, undermine public trust in social media and damage online civic culture.[46] Deep fakes could also embarrass or blackmail elected officials or individuals with access to classified information.[47] Even more worrying than disinformation spreading on social media websites such as Facebook, Instagram, or Twitter is its diffusion through encrypted messaging apps, which makes it all but impossible to track.[48] Finally, the introduction of deep fakes into classified data feeds could sow distrust in the intelligence community's conclusions.[49] If decision-makers have to assume that their intelligence collection means are compromised, this could result in their striking blindly, and potentially first, in a crisis.

[45] Tim Hwang, 'Deepfakes: A Grounded Threat Assessment', Centre for Security and Emerging Technology, July 2020, available at: https://cset.georgetown.edu/wp-content/uploads/CSET-Deepfakes-Report.pdf

[46] Cristian Vaccari and Andrew Chadwick, 'Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News', Social Media + Society, February 19, 2020, available at: https://journals.sagepub.com/doi/full/10.1177/2056305120903408

[47] For example, this technology could create non-consensual pornographic deep fakes of women in the national security infrastructure. Disinformation scholar Nina Jankowicz points out that 'Women have been enduring the trauma of deep fakes for years… [and their] participation in our representative democracy is at stake.' Analysis from Deeptrace found that in 2019 non-consensual deep fake pornography accounted for 96% of the total deep fake videos online. Source: Nina Jankowicz, 'Opinion: The threat from deepfakes isn't hypothetical. Women feel it every day.', The Washington Post, March 25, 2021, available at: https://www.washingtonpost.com/opinions/2021/03/25/threat-deepfakes-isnt-hypothetical-women-feel-it-every-day/; Henry Ajder, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen, 'The State of Deepfakes', Deeptrace, September 2019, available at: https://storage.googleapis.com/deeptrace-public/Deeptrace-the-State-of-Deepfakes-2019.pdf

[48] Nina Jankowicz, *How to Lose the Information War: Russia, Fake News, and the Future of Conflict* (London: I.B. Tauris, 2020), p. 396.

[49] Greg Allen and Taneil Chan, 'Artificial Intelligence and National Security', Belfer Center for Science and International Affairs, July 2017, available at: https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf

Satellite jamming describes interference with a satellite's signals, or the receivers of these signals. Space weather or electromagnetic 'fratricide' between nearby satellites trying to share radio-frequency spectrum can unintentionally cause jamming, however spoofing is more insidious. Satellite spoofing refers to an attempt to deceive a satellite receiver by broadcasting incorrect signals or by rebroadcasting genuine signals captured elsewhere or at a different time.[50] In other words, a spoofing attack falsifies satellite data flows to gain a strategic or tactical advantage. Though these capabilities have been around for decades, recent technological developments enable jamming across a wider spectrum of radio frequencies than before. Furthermore, seismic changes in the use of, and dependency on, space in recent years could make their use more impactful. For example, there is a growing jamming and spoofing threat towards Global Positioning Systems (GPS) for safe navigation. In December 2019, a British survey drone crashed into a house due to GPS interference.[51] The UK's Air Accidents Investigation Branch has not disclosed the source of the interference, but GPS jammers can be easily obtained on the internet, including for use as anti-drone defences.[52] Satellites are also a critical part of the nuclear infrastructure; they enable missile early warning detection, NC3, and national technical means of verification. Blinding, hoodwinking, or incapacitating such satellites could reduce a state's ability to determine if a strategic attack is underway, and reduce a state's ability to enforce arms control agreements.

These technologies are neither complex nor difficult to deploy.[53] Indeed, apps like Zao let any smartphone user create seemingly authentic deep fake videos in seconds. Satellite jamming and spoofing technologies are also ubiquitous; China, Russia, Iran, North Korea, and even non-state actors all have satellite jamming and spoofing capabilities.[54] Researchers expect that electronic counterspace technologies will 'continue to proliferate at a rapid pace in both how they are used and who is using them'.[55] Similarly, the UK Centre for Data Ethics and Innovation assesses that 'the barriers to producing deep fakes will fall and their quality will improve' raising concerns about their use by 'bad actors'.[56] This could have serious implications for crisis stability.

The technologies in this cluster are tools in the arsenal of states and non-state actors that want to operate below the threshold of overt conflict. While the effects of these technologies are easily reversed,[57] they could contribute to the environment of distrust toward institutions, including government and the media, both of which are critical players in countering disinformation.[58] The Edelman Trust Barometer have studied trust for more than twenty years. In 2021, they declared 'information bankruptcy' and recorded double-digit trust inequality figures between the 'informed public' and the 'mass population' in several countries.[59] Scholars are losing confidence that facts can prevail over disinformation, because 'falsehood diffuses significantly farther, faster, deeper, and more broadly than the truth in all categories of information'.[60] Fact-checking also presents challenges for social media companies. For example, researchers from Yale University found that the absence of warnings caused readers to perceive headlines as reliable, giving rise to an 'implied truth effect'.[61] Thus, it is not the technical characteristics of a given technology that determine its impact, but the extent to which a technology can amplify pre-existing scepticism.

Traditional concepts of crisis escalation suggest linear and somewhat predictable patterns from low-level crisis to nuclear war.[62] However, the technologies in this cluster could subvert the data flows upon which UK and its adversaries make decisions and so create

[50] INTERTANKO, 'Jamming and Spoofing of Global Navigation Satellite Systems (GNSS)', 2019, available at: https://www.maritimeglobalsecurity.org/media/1043/2019-jamming-spoofing-of-gnss.pdf

[51] Air Accidents Investigation Branch, 'AAIB investigation to DJI M600 Pro (UAS, registration n/a) 131219', June 25, 2020, available at: https://www.gov.uk/aaib-reports/aaib-investigation-to-dji-m600-pro-uas-registration-n-a-131219

[52] See, for example, Jammer4UK: https://www.jammer4uk.com/

[53] Qualitative debrief with respondents, 4 March 2021.

[54] Todd Harrison, Katilyn Johnson, Thomas G. Roberts, Tyler Way, and Makena Young, 'Space Threat Assessment 2020', Center for Strategic and International Studies, March 2020, available at: https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/200330_SpaceThreatAssessment20_WEB_FINAL1.pdf?6sNra8FsZ1LbdVj3xY867tUVu0RNHw9V

[55] Ibid.

[56] Centre for Data Ethics and Innovation, 'Snapshot Paper – Deepfakes and Audiovisual Disinformation', September 12, 2019, available at: https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation

[57] Qualitative debrief with respondents, 4 March 2021.

[58] Elizabeth Seger, Shahar Avin, Gavin Pearson, Mark Briers, Seán Ó Heigeartaigh, and Helena Bacon, 'Tackling threats to informed decision-making in democratic societies', The Alan Turing Institute, October 14, 2020, available at: https://www.turing.ac.uk/sites/default/files/2020-10/epistemic-security-report_final.pdf

[59] Edelman, 'Edelman Trust Barometer 2021', available at: https://www.edelman.com/sites/g/files/aatuss191/files/2021-01/2021-edelman-trust-barometer.pdf

[60] Soroush Vosoughi, Deb Roy, and Sinan Aral, 'The spread of true and false news online', Science, 359:6380 (2018), available at: https://science.sciencemag.org/content/359/6380/1146

[61] Gordon Pennycook, Adam Bear, Evan Collins, David G. Rand, 'The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories without Warnings', Management Science, August 7, 2019, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3035384.

[62] Herman Kahn, *On Escalation: Metaphors and Scenarios* (New York: Praeger, 1965).

escalating pathways that are less predictable. Rebecca Hersman, for example, suggests that alternative and less predictable escalatory pathways are likely, and that crisis escalation may instead follow a 'wormhole' dynamic.[63] These technologies could create 'holes' in the fabric of deterrence through which competing states could inadvertently enter and suddenly traverse between sub-conventional and strategic levels of conflict in accelerated and non-linear ways.[64]

The willingness of decision-makers to escalate a crisis will depend, in large part, on the information, cognition, and perception of what is at stake.[65] Nuclear deterrence depends on having access to information about one's own forces, doctrine, capabilities, objectives, and intent, as well as those of the adversary. This information helps decision-makers understand asymmetries in the strengths and vulnerabilities of the opposing powers, and to guide selection of appropriate strategies for influencing the adversary's behaviour in a favourable direction. This could happen, for example, by changing the cost-benefit analysis of different courses of action through a mix of coercion and inducement. While deception and disinformation have always been a feature of military competition, scholars and policymakers recognise the need for some level of transparency around doctrine and force readiness levels to build mutual trust and avoid unintended escalation. These new technologies could upend the progress made in this regard over the last five decades. Furthermore, today's decision-makers are subject to an accelerating news cycle and face political pressure to respond to events before they can be fact-checked.[66] For these reasons, these high impact and high feasibility technologies could distort the information landscape in such a way that could exacerbate an ongoing crisis.

Risk reduction measures for Cluster 1

Analysis suggested three possible measures to address the potential effects of technologies that distort.

A holistic approach to combating deep fakes. Countering deep fakes needs to be a multidisciplinary effort; it is not only a technical problem but also a psychological, human, and journalistic problem. This paper proposes three ways to combat deep fakes: detection, legislation, and education. At present, the quality and quantity of the data for deep fake detection is poor. Researchers, governments, and media companies sometimes do not have the data to analyse misinformation cases and need more data of a better quality to train deep fake detection tools.[67] Furthermore, there are more resources dedicated to audio and video manipulation than to detection.[68] Alex Engler from the Brookings Institution describes this as a 'perpetual cat and mouse game' and anticipates that automated deep fake detection is likely to become more difficult in the near future, as it becomes easier to generate fake digital content.[69] Key actors in the private sector are coming together to challenge technologies that damage public trust in video and audio. One example of this is the Coalition for Content Provenance and Authenticity (C2PA), founded by Adobe, Arm, the British Broadcasting Corporation, Intel, Microsoft, and Truepic. C2PA focuses on the development of open, global technical standards to channel content provenance efforts, including specifications and standards for social and media platforms. Future initiatives should also draw on the expertise of existing open-source intelligence and media forensics communities, as well as systems thinkers, cognitive scientists and affected communities. Although governments lag behind the private sector in terms of technical expertise, infrastructure, and their understanding of new technologies,[70] they are able to suppress audio and visual disinformation through legislation. The UK could design legislation to prohibit certain uses of a digital replica of a person. Deep fake pornography is a clear example of a harmful use of deep fake technology, which could result in sanctions such as imprisonment. This would align the UK's efforts to combat disinformation with the proposed bills in the US Congress that attempt to prohibit malicious deep fakes.[71] As regards education, governments should invest resources in educating platform users and the public on how to detect misinformation and increasing information literacy

[63] Rebecca Hersman, 'Wormhole Escalation in the New Nuclear Age', Texas National Security Review, 3:3 (2020), available at: http://dx.doi.org/10.26153/tsw/10220.

[64] Ibid.

[65] Kristin Ven Bruusgaard and Jaclyn A. Kerr, 'Crisis Stability and the Impact of the Information Ecosystem', Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict, March 15, 2020, available at: https://www.hoover.org/sites/default/files/research/docs/trinkunas_threetweetstomidnight_137-158_ch.7.pdf

[66] Ibid.

[67] Sam Gregory and Eric French, 'How do we work together to detect AI-manipulated media?' Witness Media Lab, available at: https://lab.witness.org/projects/osint-digital-forensics/

[68] Ibid.

[69] Alex Engler, 'Fighting deepfakes when detection fails', Brookings, November 14, 2019, available at: https://www.brookings.edu/research/fighting-deepfakes-when-detection-fails/

[70] Katarina Kertysova, 'Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered', Security and Human Rights, 29:1-4 (2018), available at: https://brill.com/view/journals/shrs/29/1-4/article-p55_55.xml?language=en

[71] Karen Hao, 'Deepfakes have got Congress panicking. This is what it needs to do.', MIT Technology Review, June 12, 2019, available at: https://www.technologyreview.com/2019/06/12/134977/deepfakes-ai-congress-politics-election-facebook-social/

in general.[72] For example, the UK's Essential Digital Skills Framework should include the skills necessary to identify dis- and misinformation. Public-private coalitions for content provenance, legislation prohibiting harmful deep fakes, and investing in education are three policy avenues for combating deep fakes that are applicable beyond the UK.

Protection of space-based assets linked to early warning or communications. There are two ways that space-faring actors should increase the protection of space-based assets: by protecting individual satellites and through holistic resilience approaches. These two approaches should complement each other, rather than space-faring nations viewing them as mutually exclusive. On the one hand, states and commercial actors alike should be protecting assets from known and future threats using the 'security-by-design' approach.[73] The UK government should ensure that the manufacturers of smallsats take into account the vulnerabilities of information-based systems, beginning at the early design stages.[74] But protecting individual smallsats against attack may not always be achievable, given cost constraints, the lack of norms surrounding responsible space behaviours, and enduring size, weight, and power (SWaP) requirements. So, the UK government should also focus on building resilience and mission assurance across the entire space architecture. Governments can pursue a holistic resilience approach by co-hosting payloads on multiple satellites, using larger constellations, using a range of sensors, using a range of orbits, and/or working with allies, partners, and industry to provide access. It is worth noting that whilst redundancy helps to increase security, it may also decrease security by expanding the attack surface. Put differently, increasing the number of satellites and systems in space also increases the number of vulnerabilities for adversaries to exploit. The UK and other space-faring nations should use the security-by-design and holistic resilience approaches to complement each other.

Multilateral approaches to emerging space threats and encouraging responsible space behaviours. On the multilateral level, the UK should maintain its strong international alliances and strategic defence partnerships, such as the Combined Space Operations Center (CSpOC). The CSpOC was restructured in 2018 to improve coordination between the United States, UK, and key allies, as well as between commercial and civil space organisations, to outpace emerging and advancing space threats. Furthermore, the UK is currently championing an effort to define responsible behaviours in space at the UN General Assembly (UNGA), which aims to 'increase trust and confidence between countries operating in space to prevent an arms race or a conflict that could have catastrophic consequences'.[75] This could be an appropriate forum through which to develop more effective and viable global instruments that limit the potentially dangerous consequences of electronic and cyber interference with satellites.

Cluster 2 – Compress

Experts assessed technologies in this cluster as higher impact and lower feasibility of implementation. As the largest technology cluster, it displays the greatest amount of variance, and is therefore difficult to describe in a homogenous way. This cluster contains novel means of delivering effects and new vectors of attack (ie methods that adversaries use to breach defences). The five technologies in this cluster are AI-powered cyber operations, swarm robotics, kinetic anti-satellite (ASAT) capabilities, satellites for Rendezvous and Proximity Operations (RPO) in space, and hypersonic missiles. They all impact the speed of conflict and could compress decision-making timelines. Many, but not all, are kinetic capabilities, and many may be overhyped in academic and media discourse.[76] In particular, experts pointed to the 'hype' surrounding hypersonic missiles, which material scientist Cameron L. Tracy and physicist David Wright suggest are merely 'an old technology with a massive price tag and few meaningful advantages over existing ballistic missiles'.[77] This section will describe each technology before suggesting cluster-specific recommendations.

The application of AI to offensive cyber operations provides militaries with more efficient and more effective tools for carrying out (or, in a defensive mode, containing) attacks that occur at machine speeds. In the short term, this could include making cyberattacks

[72] Elizabeth Seger, Shahar Avin, Gavin Pearson, Mark Briers, Seán Ó Heigeartaigh, and Helena Bacon, 'Tackling threats to informed decision-making in democratic societies', The Alan Turing Institute, October 14, 2020, available at: https://www.turing.ac.uk/sites/default/files/2020-10/epistemic-security-report_final.pdf

[73] Beyza Unal, 'Cybersecurity of NATO's Space-based Strategic Assets', Chatham House, July 2019, available at: https://www.chathamhouse.org/sites/default/files/2019-06-27-Space-Cybersecurity-2.pdf

[74] Ibid.

[75] HM Government, 'UK push for landmark UN resolution to agree responsible behaviour in space', August 26, 2020, available at: https://www.gov.uk/government/news/uk-push-for-landmark-un-resolution-to-agree-responsible-behaviour-in-space; Relevant submissions to the UNODA from Member States and NGOs can be found here: https://www.un.org/disarmament/topics/outerspace-sg-report-outer-space-2021

[76] Qualitative debrief with respondents, 4 March 2021.

[77] Cameron L. Tracy and David Wright, 'Don't Believe the Hype About Hypersonic Missiles', IEEE Spectrum, February 5, 2021, available at: https://spectrum.ieee.org/tech-talk/aerospace/military/hypersonic-missiles-are-being-hyped

more efficient, such as by creating better spear-phishing campaigns whereby hyper-realistic machine-written emails mimic a colleague based on their public profile or past correspondence.[78] In the medium term, this could mean machines sifting through large amounts of data in search of potential vulnerabilities, such as identifying access points and generating new malware code.[79] It is not possible to separate the offensive and defensive applications of AI-powered cyber operations; they are two sides of the same coin.[80] AI can also learn behaviour patterns for a given user environment. When it identifies abnormal or suspicious behaviour, it can alert human analysts or automate defensive actions.[81] Experts evaluated AI-powered cyber operations with a relatively lower feasibility score, highlighting the ability for human barriers, such as training requirements, and ethical/legal barriers, such as International Humanitarian Law and law of armed conflict, to limit the use of this technology by the UK.[82]

Another AI-enabled technology in this cluster is swarm robotics, which are large groups of robots that operate autonomously and coordinate their behaviour in a decentralised manner.[83] The defining feature of swarming robots is collective behaviour.[84] States and non-state actors could use robot swarms for ISR missions; perimeter surveillance and protection; distributed attacks; overwhelming enemy or missile air defences; force protection; deception; search and rescue operations; countering other swarms; and the variety of dull, dirty, and dangerous tasks already performed by individual unmanned systems.[85] The underlying algorithms rely on the premise that complex macro-level behaviours can emerge from simple local interactions between agents.[86] The strategy of exploring the available options, deciding which option to take, and communicating this to neighbouring robots compresses the amount of time available to human decision-makers.[87]

Two types of counterspace capabilities fall into this technology cluster: Rendezvous and Proximity Operations (RPO) and kinetic anti-satellite (ASAT) missiles. Neither of these are 'new' technologies, but their use against other satellites is concerning for the growing number of state and non-state actors who are reliant upon a sustainable space environment. RPO describes a satellite approaching, or even making contact with, another, for example, for the purposes of on-orbit refuelling or repair. RPO applications are vast and span commercial, civil and military uses. Whilst they are promising in their ability to contribute to the sustainability of space, for example, through active debris removal, concerns arise from their co-orbital ASAT capability.[88] In this context, a co-orbital ASAT capability describes a satellite denying or degrading another satellite from the same orbit. The Secure World Foundation (SWF) and the Center for Strategic and International Studies (CSIS) report that there is evidence to suggest that states, including Russia, have developed co-orbital ASAT capabilities.[89]

Kinetic ASAT capabilities are more straightforward in their impact, compared with co-orbital ASAT capabilities. Launched from earth, the former use

[78] Dave Palmer, 'AI will supercharge spear phishing', Darktrace Blog, January 9, 2017, available at: https://www.darktrace.com/en/blog/ai-will-supercharge-spear-phishing

[79] Pieter Arntz, Wendy Zamora, Jérôme Segura, and Adam Kujawa, 'When artificial intelligence goes awry: separating science fiction from fact', Malwarebytes Labs, 2019, available at: https://resources.malwarebytes.com/files/2019/06/Labs-Report-AI-gone-awry.pdf

[80] See, for example, the U.S. Department of Defence, Cyber Strategy, 2018, available at: https://media.defense.gov/2018/Sep/18/2002041658/-1/-1/1/CYBER_STRATEGY_SUMMARY_FINAL.PDF

[81] Darrell M. West and John R. Allen, 'How artificial intelligence is transforming the world', Brookings, April 24, 2018, available at: https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world

[82] Notwithstanding the expert-assessed lower feasibility score, the UK is a global leader in cyberspace (third globally, per the National Cyber Power Index 2020) and is home to some of the world's top AI companies, which makes its use of AI for cyber operations seem probable. Responses from experts indicate that the UK is likely to utilise AI for cyber operations in a more targeted and restrained way.

[83] Maaike Verbruggen, 'Drone swarms: coming (sometime) to a war near you. Just not today.', Bulletin of the Atomic Scientists, February 3, 2021, available at: https://thebulletin.org/2021/02/drone-swarms-coming-sometime-to-a-war-near-you-just-not-today

[84] The collective behaviour of a swarm is often bio-inspired, meaning that individual robots cooperatively solve problems by mimicking structures and behaviours similar to those observed in natural systems, such as bees, birds, or fish. Source: Melanie Schranz, Martina Umlauft, Micha Sende, and Wilfried Elmenreich, 'Swarm Robotic Behaviors and Current Applications', Frontiers in Robotics and AI, April 2, 2020, available at: https://www.frontiersin.org/articles/10.3389/frobt.2020.00036/full

[85] Merel Ekelhof and Giacomo Persi Paoli, 'Swarm Robotics: Technical and Operational Overview of the Next Generation of Autonomous Systems', UNIDIR, 2020, available at: https://www.unidir.org/publication/swarm-robotics-technical-and-operational-overview-next-generation-autonomous-systems

[86] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz, Swarm Intelligence: From Natural to Artificial Systems (Oxford: Oxford University Press, 1999).

[87] Gabriele Valentini, Heiko Hamann, and Marco Dorigo, 'Efficient Decision-Making in a Self-Organizing Robot Swarm: On the Speed Versus Accuracy Trade-Off', Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), May 2015, available at: https://www.researchgate.net/publication/272495967_Efficient_Decision-Making_in_a_Self-Organizing_Robot_Swarm_On_the_Speed_Versus_Accuracy_Trade-Off

[88] Anuradha Damale, 'Rendezvous Proximity Operations: Not operating in isolation', European Leadership Network, August 12, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/rendezvous-proximity-operations-not-operating-in-isolation

[89] Secure World Foundation, 'Global Counterspace Capabilities', available at: https://swfound.org/counterspace/; Todd Harrison, Katilyn Johnson, Thomas G. Roberts, Tyler Way, and Makena Young, 'Space Threat Assessment 2020', Center for Strategic and International Studies, March 2020, available at: https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/200330_SpaceThreatAssessment20_WEB_FINAL1.pdf?6sNra8FsZ1LbdVj3xY867tUVu0RNHw9V

physical force to damage or destroy target satellites. The US, Russia, China, and India have all successfully tested direct ascent ASAT capabilities, while Israel and others claim to have the necessary technology.[90] Although this capability dates to the early space age (1959), there are no records of hostile use; however, the development, testing and demonstration of such capabilities by all the aforementioned states is accelerating.[91] In the case of both kinetic and co-orbital ASATs, the damage to the target satellite is often irreversible, as is the potential increase of space debris in orbit.[92] This is likely to impact terrestrial inter-state tensions. Finally, space-based systems play an important role in most nuclear states' NC3 networks, including early-warning sensors.[93] Concerning the impact of an ASAT attack on crisis stability, Karl Mueller writes that, 'Under conditions of real or perceived first-strike advantage... decision-making timelines are likely to be very compressed.'[94] He suggests that this can contribute to a 'witch's brew of pathological effects' for nuclear decision-makers, including limited opportunities for communication and signalling between adversaries, as well as constrained collection and analysis of information.[95]

The fifth technology in this cluster is hypersonic missiles. Hypersonic missiles travel at least five times the speed of sound (ie greater than Mach 5 or approximately 6,174 km per hour), often with the ability to manoeuvre during flight. NATO claims that hypersonic missiles can bypass existing early warning systems and expects them to be extremely stealthy.[96] Ambiguity surrounding whether or not a weapon is nuclear armed also creates some concern about these systems. At present, the United States, China, and Russia have hypersonic missiles, and research programmes are underway in Japan, India, North Korea, South Korea, France, Australia, and the United Kingdom. 'Hypersonic missiles' refers to two sub-classes of weapons: hypersonic glide vehicles (HGV) are launched from a rocket before gliding to a target, whereas hypersonic cruise missiles (HCM) are powered by high-speed, air-breathing engines during flight. Most experts reported that they had HGVs in mind when they evaluated this technology. Experts assessed hypersonics as the most likely technology (of all 10 on the shortlist) to deliver, or enable the delivery of, a disarming first strike and/or to elicit a nuclear response. They were also assessed as reducing decision-making time to the greatest extent of all the technologies.

Given the heterogeneity of this technology cluster, there is no single story about the impact that these technologies will have on crisis stability. However, experts assessed that these five technologies could accelerate an ongoing crisis by compressing decision-making time or by enabling a disarming first strike. This could eliminate an adversary's ability to retaliate with nuclear weapons and thereby create a 'use it or lose it' situation that would incentivise an adversary to strike first in a crisis.

Risk reduction measures for Cluster 2

Analysis suggested several possible measures to address the potential effects of technologies that compress decision-making time.

'Traditional' arms control. Formal treaties that are enforceable, legally-binding agreements, would be one of the best options for regulating technologies that compress decision-making time. Scholars Madeline Zutt and Michal Onderco from Erasmus University Rotterdam write that, 'Regulating a technology such as hypersonic missiles is comparatively more straightforward because it can be done using traditional arms control tools.'[97] This could take the form of a multinational ban on exports of complete hypersonic delivery vehicles and case-by-case export reviews for hypersonic missile subsystems. This should encompass hypersonic fuels and flight controls, supersonic combustion ramjet engines, warheads, etc.[98] This would only require

[90] Dan Williams, 'Israel says Arrow 3 missile shield aces test, hitting target in space', Reuters, December 10, 2015, available at: https://www.reuters.com/article/us-arms-israel-arrow-trial-idUSKBN0TT0HU20151210

[91] Ben Skinner, 'Ground-Based Weapons: Kinetic Antisatellite Weapons', Space Security Index, April 2020, available at: https://spacesecurityindex.org/2020/04/ground-based-weapons-kinetic-antisatellite-weapons

[92] Space debris, combined with a higher overall number of operational objects in space, could result in the Kessler syndrome which describes the cascading effect of a scenario wherein the high density of objects causes collisions, generating debris that in turn increases the likelihood of future collisions (referred to as 'collisional cascading'), and ultimately rendering parts or entire orbits unusable for further space activity.

[93] Timothy Wright, 'Do ASATs mean less security in space?', IISS, March 17, 2020, available at: https://www.iiss.org/blogs/military-balance/2020/03/india-anti-satellite-weapon-space-security

[94] Karl Mueller, 'The Absolute Weapon and the Ultimate High Ground: Why Nuclear Deterrence and Space Deterrence Are Strikingly Similar – Yet Profoundly Different', in Anti-satellite Weapons, Deterrence and Sino-American Space Relations, September 2013, available at: https://www.stimson.org/wp-content/files/file-attachments/Anti-satellite%20Weapons%20-The%20Stimson%20Center.pdf

[95] Ibid.

[96] NATO Science & Technology Organization, 'Science & Technology Trends 2020-2040: Exploring the S&T Edge', March 2020, available at: https://www.nato.int/nato_static_fl2014/assets/pdf/2020/4/pdf/190422-ST_Tech_Trends_Report_2020-2040.pdf

[97] Madeline Zutt and Michal Onderco, 'How emerging technologies impact the future of nuclear risk and arms control', European Leadership Network, September 1, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/how-emerging-technologies-impact-the-future-of-nuclear-risk-and-arms-control

[98] Christopher A. Bidwell and Bruce W. MacDonald, 'Emerging Disruptive Technologies and Their Potential Threat to Strategic Stability and National Security', Federation of American Scientists, September 2018, available at: https://fas.org/wp-content/uploads/media/FAS-Emerging-Technologies-Report.pdf

participation from the nations who possess hypersonic missiles: the United States, Russia and China.[99] Because a number of regimes for technology export controls currently exist, negotiators can extend the lessons learned from a substantial body of experience to the regulation of hypersonic missiles. Critically though, this is not the case for all technologies in this cluster, many of which cannot be quantified or tracked in the same way that tangible assets, such as missiles, can be counted.

Building norms of responsible behaviours. Policymakers cannot limit their arms control ambitions to legally-binding agreements because there are many situations where the technical complexity of a new technology and the political factors that frame the debate may make this particularly challenging. Even when they are possible, treaties may take years to negotiate, which limits their ability to control the developmental trajectory of a rapidly emerging technology. For example, when it comes to military AI and counterspace capabilities, experts are increasingly moving away from regulating capabilities and towards regulating behaviours. This paper has already mentioned the UNGA resolution on 'Reducing Space Threats through Norms, Rules and Principles of Responsible Behaviours', initially proposed by the UK. Likewise, NATO has initiated a process that encourages its members to agree on responsible governance principles for new technologies.[100] And the Czech Republic, Finland, Germany, the Netherlands, and Sweden are calling on the European Union to start a process on the responsible military use of new technologies.[101] Short of a legally-binding treaty, 'soft' instruments like norms, codes of conduct and agreed standards of behaviour might provide a basis for building more formal treaties at a later date.

Strategic cyber NFU policy. According to General Sir Patrick Sanders of Strategic Command, the objective of the UK's National Cyber Force (NCF) is 'to defend the UK in cyberspace'.[102] It does so by disrupting the activities of those groups and nations it deems to be sufficient threats, including hostile states, terrorists, and criminals.[103] This is similar to the United States' strategy of 'defending forward' in cyberspace, which also blurs the line between defensive and offensive cyber operations.[104] 'Defending forward' entails proactive observing, pursuing, and countering of adversary operations. Public acknowledgement of the existence of these programmes reflects an effort on the part of the UK and US governments to increase transparency in this domain. Even so, given the nature of the capabilities, there is still an understandable lack of transparency about the work of the NCF.[105] Ambiguity around what the NCF can do in the name of defence could increase misperceptions in a crisis. To combat such misconceptions, scholars such as Jacquelyn Schneider suggest that states should extend declaratory policies of no first use (NFU) into the cyber domain.[106] A strategic cyber NFU could create norms of restraint to decrease the incentives for developing and launching destabilising cyber-attacks. For example, a strategic cyber NFU could discourage cyber-attacks against critical national infrastructure or nuclear forces. It could also minimise the chance of crisis escalation at the hands of cyber operations that are both defensive and offensive. This could shore up strategic deterrence for nuclear possessors like the UK and US.

Nationally assured space situational awareness (SSA). Recent trends in the use of space systems by the military and the commercialisation of the sector have created new dependencies, risks, and vulnerabilities for space-faring actors. This presents both national security and resilience concerns. An effective space situational awareness capability and trusted process, with global coverage and involvement of NWS, NNWS, and private actors will be critical to mitigate the risks of collision and debris creation, as well as the risks to satellite owners, operators and end-users who are reliant on space-enabled services. Operational SSA provides an overview of the environment; as of 15 April 2021, there are an estimated 34,000 objects greater than 10 centimetres in orbit, 900,000 objects between 1 and 10 centimetres, and 128 million objects between 1 millimetre and 1 centimetre.[107]

[99] Richard H. Speier, George Nacouzi, Carrie Lee, and Richard M. Moore, 'Hypersonic Missile Nonproliferation: Hindering the Spread of a New Class of Weapons', RAND Corporation, 2017, available at: https://www.rand.org/pubs/research_reports/RR2137.html

[100] Edward Hunter Christie, 'Artificial Intelligence at NATO: dynamic adoption, responsible use', NATO Review, November 24, 2020, available at: https://www.nato.int/docu/review/articles/2020/11/24/artificial-intelligence-at-nato-dynamic-adoption-responsible-use/index.html

[101] 'Minister's Declaration at the occasion of the Conference', Capturing Technology, Rethinking Arms Control, November 6, 2020, available at: https://rethinkingarmscontrol.de/wp-content/uploads/2020/11/Ministerial-Declaration-RAC2020.pdf

[102] Ministry of Defence, Foreign, Commonwealth & Development Office, Government Communications Headquarters, Defence Science and Technology Laboratory, and The Rt Hon Ben Wallace MP, 'National Cyber Force Transforms country's cyber capabilities to protect UK', November 19, 2020, available at: https://www.gov.uk/government/news/national-cyber-force-transforms-countrys-cyber-capabilities-to-protect-uk

[103] Ibid.

[104] The U.S. Department of Defence, Cyber Strategy, 2018, available at: https://media.defense.gov/2018/Sep/18/2002041658/-1/-1/1/CYBER_STRATEGY_SUMMARY_FINAL.PDF

[105] Matt Burgess, 'The UK created a secretive, elite hacking force. Here's what it does', WIRED, November 20, 2020, available at: https://www.wired.co.uk/article/national-cyber-force-uk-defence-gchq

[106] Jacquelyn Schneider, 'A Strategic Cyber No-First-Use Policy? Addressing the US Cyber Strategy Problem', *The Washington Quarterly,* 43:2 (2020), available at: https://www.tandfonline.com/doi/abs/10.1080/0163660X.2020.1770970?journalCode=rwaq20

[107] European Space Agency, 'Space debris by the numbers', April 15, 2021, available at: https://www.esa.int/Safety_Security/Space_Debris/Space_debris_by_the_numbers

This is important because small fragments of space debris can have disproportionately large impacts on space vehicles and space-based infrastructure, due to their velocity. However, SSA is not sufficiently advanced to monitor behaviours in space, meaning that it cannot ensure the content and appropriateness of RPO, for example.[108] Furthermore, there is currently no single provider for global SSA oversight, which introduces the possibility of conflicting information.[109] Given the increasing number of space operators (both public and private) and congestion in space, the UK should develop a nationally assured SSA capability and trusted process. This could be as simple as two strategically placed radars, one in the Falklands and the second in Northern Scotland, which would provide significant support to the ability to monitor objects in low Earth orbit (LEO).[110] The UK should develop this as a civil programme because it is easier to share civilian data with the military than the other way around. This could augment the Allied Space Surveillance Network, which is global in scope, but cannot share data behind the 'military firewall' and is optimised for missile warning, not SSA.[111] A nationally assured SSA capability will be critical to the UK's mission of 'enhanc[ing] space sustainability and maintain[ing] the UK space industry as a global leader'.[112] The development of this capability could be a logical technical solution for other NWS and NNWS to address the issues posed by the uncertainty surrounding counterspace activities. Minimising uncertainty, for example, in attribution, would also minimise the chance of miscalculation and increase crisis stability.

Cluster 3 – Thwart

The algorithm assigned only one technology, directed energy weapons (DEWs), to this cluster. Experts assessed this technology as lower impact and lower feasibility of implementation. It is somewhat difficult to generalise the characteristics of a cluster of one. However, several features set DEWs apart from the nine other technologies in the shortlist, which will require distinct risk reduction measures. Experts evaluated that the primary impact of this cluster on crisis stability is its ability to credibly prevent or blunt a nuclear attack. Experts characterised DEWs as currently being 'more defensive than offensive' and 'not adversarial or offensive'.[113] But augmenting defence may also be destabilising if it has the intended or ancillary effect of diminishing a country's second-strike response to a first strike.[114] For example, both Russia and China argue that US missile defences undermine their nuclear second-strike capabilities and use this as a justification to modernise their nuclear arsenals.

A DEW damages or incapacitates its target with highly focused energy, including laser, microwaves, and particle beams. DEWs have been the subject of speculation for strategic weapons applications for decades. For example, the 1980s US Strategic Defense Initiative, nicknamed 'Star Wars', envisioned the potential use of DEWs for missile defence. However they may be considered as an emerging technology due to recent advances, particularly solid-state lasers. If matured, DEWs may provide several advantages over traditional kinetic weapons due to their speed-of-light delivery, precision engagement, scalable effects, and low cost per engagement. Furthermore, DEWs are a covert capability without visual signs, making them difficult to detect. For these reasons, DEWs are likely to become an important part of the missile defence systems of the future. Christopher Bidwell and Bruce MacDonald explain that 'applications of high-energy lasers for boost-phase missile defence from aerial platforms – either unmanned aerial vehicles (UAVs) or aircraft – could be a serious challenge to fixed-base, highly 'MIRVed' (multiple independently targetable re-entry vehicle) ICBMs'.[115] Notwithstanding these strengths, there are also various drawbacks to DEWs, including their reliance on an unobstructed path, their high energy demands, their poor performance in bad weather, their potential need to have the beam on target for a period of time

[108] Anuradha Damale, 'Rendezvous Proximity Operations: Not operating in isolation', European Leadership Network, August 12, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/rendezvous-proximity-operations-not-operating-in-isolation

[109] Bruce McClintock, Katie Feistel, Douglas C. Ligor, and Kathryn O'Connor. 'Responsible Space Behaviour for the New Space Era: Preserving the Province of Humanity', RAND Corporation, 2021, available at: https://front.un-arm.org/wp-content/uploads/2021/04/rand-pea887-2.pdf

[110] Interview with expert, May 7, 2021.

[111] Joint Air Power Competence Centre, 'Command and Control of a Multinational Space Surveillance and Tracking Network', June 2019, available at: https://www.japcc.org/wp-content/uploads/JAPCC_C2SST_2019_screen.pdf

[112] UK Space Agency, Ministry of Defence, and Department for Business, Energy & Industrial Strategy, ''Government backs UK companies tackling dangerous 'space junk'', September 16, 2020, available at: https://www.gov.uk/government/news/government-backs-uk-companies-tackling-dangerous-space-junk

[113] Qualitative feedback from experts upon reverting the scoring exercise.

[114] Christopher F. Chyba, 'New Technologies and Strategic Stability', Daedalus, 149:2 (2020), available at: https://www.jstor.org/stable/10.2307/48591318

[115] Christopher A. Bidwell and Bruce W. MacDonald, 'Emerging Disruptive Technologies and Their Potential Threat to Strategic Stability and National Security', Federation of American Scientists, September 2018, available at: https://fas.org/wp-content/uploads/media/FAS-Emerging-Technologies-Report.pdf

to achieve effect, and their enduring difficulties with slow-to-mature battery technologies.

The United States, Russia, United Kingdom, Israel, China, and Japan have all developed DEWs.[116] In practice, functions range from anti-drone laser weapons such as China's Silent Hunter, the United States' High Energy Laser with Integrated Optical-dazzler and Surveillance (HELIOS) system, and the UK's ship-borne laser weapon 'Dragonfire'. The next generation of ground-based DEWs are likely to be so highly powered and highly focused that they can reach beyond LEO and pose a threat to GPS satellites in medium Earth orbit.[117] Furthermore, the US Missile Defence Agency submitted a budget request to Congress in March 2019 to 'design, develop, and conduct a feasibility demonstration for a space-based, directed energy intercept layer'.[118] Critics argue that deploying space-based DEWs would be 'technologically challenging', 'prohibitively expensive', and potentially destabilising.[119]

In the future, DEWs might form a valuable part of a layered defensive system-of-systems, including missile interceptors and other capabilities. On the one hand, this could enhance stability by deterring an attack, for example, by deterrence through denial or by altering the risk/benefit calculation of the attacker. On the other hand, there is a chance that another state could see investment in DEWs as undermining the credibility of its second-strike capabilities, contributing to a 'use it or lose it' dynamic.

Risk reduction measures for Cluster 3

Analysis suggested two possible measures to address the potential risks of DEWs during a crisis:

Limiting the number of DEWs that can be deployed. DEWs should be treated in tandem with concerns about ballistic missile defence (BMD) as potentially undermining crisis stability by reducing the likelihood of a successful second strike. One policy option could be to limit the number of such systems that could be deployed, including both ground and air vehicle-based DEWs. Unfortunately, this would run up against US policy of resisting limitations on BMD.

Norms against placing DEWs in space. Space has always been militarised, but states could take steps to avoid further weaponization of space that could lead to crisis instability. The 1967 UN Treaty on Outer Space remains the authoritative overarching set of international obligations for all space activities. It can be argued that the pace of technological developments requires more up-to-date norms to regulate activities in space more effectively. A ban on space-based DEWs or norms regulating their use could be a long-term objective for the UNGA as part of the draft resolution on 'Reducing Space Threats Through Norms, Rules and Principles of Responsible Behaviours'.

Cluster 4 – Illuminate

Experts assessed these technologies as lower impact and higher feasibility of implementation. This cluster includes two technologies: smallsats and AI for ISR missions. Experts agreed that this cluster exhibits a high degree of homogeneity and conceptual clarity.[120] These are the only technologies that were identified by experts as capable of both strengthening *and* eroding NC3, as well as both increasing *and* reducing decision-making time and situational awareness. In this way, they simultaneously represent an opportunity and a challenge to crisis stability. While these technologies can potentially provide more accurate and comprehensive data flows to decision-makers and increase situational awareness, interfering or tampering with them could destabilise a crisis.

Smallsats are distinct from the space technologies in other clusters because they are not counterspace technologies. Instead, they enable a range of civilian services in the finance, transportation, and crisis management sectors that rely on positioning, navigation, timing, meteorological services, telecommunications, and Earth observation. Critically, dense constellations of smallsats in LEO also provide ISR capabilities, which are continuous, collectively survivable, and available on-demand for tactical warfighting applications.[121] For example, the UK's Carbonite satellites record high-definition, full-colour video with a resolution

[116] GlobalData Thematic Research, 'Directed Energy Weapons: Timeline', Army Technology, August 11, 2020, available at: https://www.army-technology.com/comment/directed-energy-weapons-laser

[117] U.S. Defense Intelligence Agency, 'Challenges to Security in Space', February 11, 2019, available at: https://www.dia.mil/Portals/27/Documents/News/Military%20Power%20Publications/Space_Threat_V14_020119_sm.pdf

[118] Kingston Reif, 'U.S. Seeks New Space-Based Capabilities', Arms Control Association, April 2019, available at: https://www.armscontrol.org/act/2019-04/news/us-seeks-new-space-based-capabilities

[119] Ibid.

[120] Qualitative debrief with respondents, 4 March 2021.

[121] Chris Simi, 'Small Satellite Sensors', DARPA, available at: https://www.darpa.mil/program/small-satellite-sensors

[122] Surrey Satellite Technology Limited (SSTL), 'Earth Observation Spacecraft', available at: https://www.sstl.co.uk/what-we-do/earth-observation-spacecraft

of one metre. Some of its applications include pattern of life assessments, humanitarian and disaster management, situational awareness, national security, and infrastructure and asset monitoring.[122] The bottom line is that smallsats could augment situational awareness in a crisis. However, there are some important practical limitations of using smallsats for ISR, including the trade-offs associated with reduced size, weight, and power (SWaP) requirements and the fact that individual satellites will only be over the target area for a moment. There are some enduring advantages to using large traditional ISR satellites in other orbits or using platforms such as high-altitude pseudo satellites.[123] Furthermore, the main distinguishing factor of smallsat launch is that tens or hundreds of satellites can be launched using the same rocket, or as secondary payloads alongside the launch of a larger satellite.[124] This will increase orbital congestion, particularly in LEO, and increase the risk of collision, as discussed in Cluster 1.

The second technology in this cluster refers to the use of AI for ISR. Existing sensors collect far too much data to sift through manually, especially when operators must make critical decisions quickly, such as offensive counter air and defensive counter air missions. The approach helps to synthesise oceans of data into actionable intelligence and accurate targeting information at speed and on a greater scale. This enables the faster and more accurate collection and synthesisation of data and facilitates more informed command and control decisions. Although the incorporation of ML and autonomous systems can lessen the data searching, processing, and analysis burden for human command, the inclusion of technical elements contribute to system complexity and so create a new source for errors, biases or vulnerabilities hidden from operators.[125] By creating both new opportunities and new vulnerabilities, AI for ISR could both strengthen and erode NC3. Whereas the reliable functioning of AI for ISR would increase situational awareness, spoofing such systems would decrease situational awareness.

There is a duality to the technologies in in Cluster 4, which simultaneously represent an opportunity and a challenge to crisis stability. However, experts were generally optimistic that these two technologies would positively impact crisis stability, particularly if manufacturers address security concerns at the design stage. Experts assessed these technologies as very unlikely to elicit a nuclear response in and of themselves and assessed them as being more likely to strengthen NC3 and increase decision-making time and situational awareness for the user than to erode or reduce them. Experts also remarked that these technologies underpin and enable those in Cluster 1, further emphasising the two ‘faces’ of these technologies and their multipurpose nature.[126]

Risk reduction measures for Cluster 4

Analysis suggested one possible measure to address the potential effects of technologies that illuminate.

NWS commitment not to target each other’s NC3 infrastructure. The threat of attacks on NC3 have been discussed at length over the last several years, with the recognised NWS under the NPT factoring this threat into their deterrence policies.[127] An achievable risk reduction measure would be for each of the P5 states separately to commit not to target one another’s NC3 infrastructure, in the knowledge that doing so could be escalatory. This is one way that the UK, or other NWS, could achieve its stated goals of ‘seek[ing] to create dialogue among states possessing nuclear weapons… to increase understanding and reduce the risk of misinterpretation and miscalculation.’[128] In order to alleviate the concerns of NNWS, the NWS within the P5 process could reiterate this commitment at the upcoming NPT Review Conference. To do so would modernise the tradition of declaratory commitments, perhaps the most notable of which is the Regan-Gorbachev statement: ‘A nuclear war cannot be won and must never be fought.’[129]

[123] See, for example, Zephyr: https://www.airbus.com/defence/uav/zephyr.html

[124] For example, a single SpaceX Falcon 9 rocket launched 60 Starlink internet satellites into orbit on April 2, 2021. Starlink is the name of a satellite network ‘megaconstellation’ that the private spaceflight company SpaceX is developing to provide low-cost internet to remote locations. SpaceX eventually hopes to place as many as 12,000 satellites in low Earth orbit (LEO).

[125] Wilfred Wan, ‘Nuclear Risk Reduction: A framework for analysis’, UNIDIR, 2019, available at: https://unidir.org/publication/nuclear-risk-reduction-framework-analysis

[126] Qualitative debrief with respondents, 4 March 2021.

[127] Dmitry Stefanovich, ‘Russia’s Basic Principles and the Cyber-Nuclear Nexus’, European Leadership Network, July 14, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/russias-basic-principles-and-the-cyber-nuclear-nexus

[128] HM Government, ‘Global Britain in a competitive age: The Integrated Review of Security, Defence, Development and Foreign Policy’, March 2021, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age-_the_Integrated_Review_of_Security__Defence__Development_and_Foreign_Policy.pdf

[129] Ronald Reagan Presidential Library and Museum, ‘Joint Soviet-United States Statement on the Summit Meeting in Geneva’, November 21, 1985, available at: https://www.reaganlibrary.gov/archives/speech/joint-soviet-united-states-statement-summit-meeting-geneva

Risk reduction recommendations for nuclear possessors and non-possessors

This section proposes broader risk reduction measures, with relevance to states with and without nuclear weapons.

Cooperation with the private sector. As previously mentioned, many of the emerging technologies discussed in this paper originate in an ecosystem that is fundamentally different from the traditional defence industrial model, which was more top-down in nature, with a small number of sellers and a single buyer, typically the military.[130] In contrast, many of the technologies discussed in this study are already being developed in the private sector, often by multinational companies that have not traditionally worked for defence. This is a more bottom-up model. The public sector must therefore be able to communicate with the private sector about the potential harms of dual-use technologies and explain why it may be worthwhile for a wider range of defence suppliers to consider the security needs of society. This proposal should help to address dual-use commercial off-the-shelf technologies (ie goods and services that are available in the commercial marketplace). Private actors also ought to consider how their commercial products could become an unwitting part of another state's foreign policy objectives, through what Doug Britton calls the 'Cyber-Newtonian Wormhole',[131] and consider the potentially negative external costs of their research investments. To counteract this, NWS could partner with private companies that are driving innovation to develop risk reduction measures specifically in relation to nuclear weapons, such as NC3.

Bridge-building between technology and policymaking. A second, related, recommendation is that NWS should provide incentives to recruit into government from the private sector. The World Economic Forum acknowledges that bridging the divide between technology and policymaking will be critical to avoid the pitfalls of technologies, while benefiting from their promise.[132] On the supply side, this could involve changes to educational curricula to include technology ethics and human-centred design. On the demand side, there need to be more, and better paid, roles for scientists and technologists who want to work in public policy, in government agencies and legislative staffs. Furthermore, the UK should create new formats of engagement so that policymakers and scientists/technologists can come together and discuss emerging technologies and their interconnections. To use an example from another sector, the UK government acknowledges that 'to be successful, industry, science, policymakers will need to work together' to stay at the cutting edge of transport innovation.[133] Similar principles should apply to security and defence.

Bridge-building between NWS and NNWS. Analysis of these technology clusters highlights both challenges and opportunities for reducing the risks of crisis escalation. Ideally, NWS and NNWS would work together to address these issues, as they have done with similar challenges such as disarmament verification. Since the mid-2000s, the UK has positioned itself as a leader in research on verification of nuclear disarmament and has worked with prominent NNWS such as Norway to lay the technological groundwork for future arms control agreements.[134] The UK also pursues these goals via the International Partnership for Nuclear Disarmament Verification (IPNDV) and the Quad Nuclear Verification Partnership,[135] which bring together

[130] Antonio Missiroli, 'Game of drones? How new technologies alter deterrence, defence and security', NATO Review, May 5, 2020, available at: https://www.nato.int/docu/review/articles/2020/05/05/game-of-drones-how-new-technologies-affect-deterrence-defence-and-security/index.html

[131] Doug Britton, 'The CANOPY WING Vulnerability: Weaponizing the Weakness', IIoT World, October 24, 2018, available at: https://iiot-world.com/ics-security/cybersecurity/the-canopy-wing-vulnerability-weaponizing-the-weakness

[132] Bruce Schneier, 'We must bridge the gap between technology and policymaking. Our future depends on it', World Economic Forum, November 12, 2019, available at: https://www.weforum.org/agenda/2019/11/we-must-bridge-the-gap-between-technology-and-policy-our-future-depends-on-it

[133] Government Office for Science, 'A time of unprecedented change in the transport system', January 2019, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/780868/future_of_mobility_final.pdf

[134] See, for example, United Kingdom – Norway Initiative (UKNI): https://ukni.info/#:~:text=United%20Kingdom%20%2D%20Norway%20Initiative,arms%20control%20and%20disarmament%20verificationd

[135] The Quad Verification Partnership includes membership from the United Kingdom, Norway, Sweden, and the United States. For more information, see: https://quad-nvp.info

NWS and NNWS to identify challenges associated with nuclear disarmament verification and develop solutions to address those challenges. Both groupings have something to gain from these arrangements: NWS have much to learn about the effects of strategic technologies on civilian populations; NNWS gain an opportunity to build skills and capacities; and both parties build stronger partnerships. Expanding existing verification partnerships and continuing dialogue in this area would help to further increase all states' level of understanding around different types of nuclear risk globally.[136] The UK and other NWS should expand the nature of their collaboration with NNWS, such as Estonia and the Netherlands, who have specialist knowledge in cybersecurity. Such partnerships could share relevant threat information, joint analysis, and conduct coordinated investigations to attribute cyber-attacks. Such partnerships could extend beyond defence to include, for example, digital healthcare.[137]

Developing confidence-building measures to increase trust between P5 members. At the most recent full P5 meeting, in February 2020, the UK was one of the leading voices calling for increased discussion of emerging technologies and their impact on nuclear stability. Emerging technologies are now a regular agenda item for the P5 process. Ideally, these conversations will lead to the development of near-term politically-binding confidence-building measures that can help create and sustain mutual understanding and trust between P5 members. These measures could include regular dialogue, information sharing, best practice exchanges, and scientific cooperation programmes. Such agreements might resemble historical efforts to avoid misperceptions during crises, such as the 1972 Incidents at Sea Agreement, but making progress on risk reduction around these technologies will also require new thinking and mechanisms.

Using emerging technologies to support nuclear risk reduction. Identifying the potential risks and challenges posed by these technologies was a primary objective of this study. But these technologies also offer potential benefits to stability. The technology clustering exercise returned one cluster (Cluster 4) that highlighted the potential for states to integrate AI and smallsats into their tool kit for early warning, detection, and target identification. Doing so could prevent close calls and increase an actor's understanding of adversary actions.[138] There are a range of other emerging technologies that present opportunities to support compliance and verification regimes, including distributed ledger technology for nuclear materials control,[139] image recognition for verification activities,[140] metadata for geolocation,[141] AI and synthetic environments for improved military planning and wargaming,[142] among others. NWS and NNWS should seriously consider the ability of these and other emerging technologies to improve their overall capabilities in the nuclear space and reduce the risk of unintended escalation during a crisis. For example, parties to the IPNDV and the Quad should discuss technologies that enable new verification activities.

Strategic stability dialogue. One final risk reduction measure that could help to mitigate the risks of emerging technologies would be a high-level dialogue, where parties come to the table prepared to discuss strategic stability, deterrence, and nuclear risks, with a focus on risks created by emerging technologies. At the time of writing, this avenue seems plausible. On 16 April 2021, hours after signing the executive order imposing sanctions on Russia, US President Joe Biden proposed a strategic stability dialogue with his counterpart Vladimir Putin.[143] Though prompted by 'Russia's history of carrying out reckless and disruptive cyber operations',[144] the summit, to be held in Europe in Summer 2021, will address a range of critical global challenges, including nuclear risk. However, it is important to acknowledge that political will is crucial in implementing risk reduction measures, especially when it comes to maintaining dialogue over a longer period.[145]

[136] Marion Messmer, 'Strategic Risk Reduction in the European Context', BASIC, June 2020, available at: https://basicint.org/wp-content/uploads/2020/06/Strategic-Risk-Reduction-in-the-European-Context-WEB-1.pdf

[137] Ciaran Martin's speech in Tallinn, Estonia, September 14, 2017, available at: https://www.ncsc.gov.uk/speech/ciaran-martins-speech-tallinn-estonia

[138] Jessica Cox and Heather Williams, 'The Unavoidable Technology: How Artificial Intelligence Can Strengthen Nuclear Stability', The Washington Quarterly, 44:1 (2021), available at: https://www.tandfonline.com/doi/abs/10.1080/0163660X.2021.1893019

[139] See, for example, SLAFKA: https://www.stimson.org/2020/dlt-prototype-for-nuclear-materials-control

[140] Zach Dorfman, 'True detectives', Middlebury Magazine, May 23, 2018, available at: http://middleburymagazine.com/features/true-detectives

[141] Melissa Hanham, 'Metadata: MetaUseful & MetaCreepy', Bellingcat, April 24, 2015, available at: https://www.bellingcat.com/resources/how-tos/2015/04/24/metadata-metauseful-metacreepy

[142] See, for example, Athena: https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries-smarter-faster-and-more-efficient-with-artificial-intelligence

[143] Morgan Chalfant and Brett Samuels, 'Biden calls for dialogue with Russia amid raft of sanctions', The Hill, April 15, 2021, available at: https://thehill.com/homenews/administration/548550-biden-calls-for-dialogue-with-russia-after-sanctions

To ignore emerging technologies increases nuclear risks

The objective of this study was to help policymakers identify how emerging technologies might increase nuclear risks and which technologies should be the focus of multilateral efforts to reduce those risks. It offers a framework for evaluating diverse technologies in a way that makes them comparable, by grouping technologies with similar risk profiles into technology clusters. As identified by the experts surveyed in this study, the most concerning technology cluster contains technologies that could 'distort' the information space and consequently undermine trust in the quality of information, damage online civic culture, and potentially escalate a crisis in non-linear ways.

Technology clusters allow policymakers to compare different technologies in terms of common parameters. This is important because it helps states allocate resources, offers a new means for addressing emerging technologies in collaborative ways, and ultimately contributes to more comprehensive cross-domain risk reduction. Finally, this framework provides a new way for NWS and NNWS to advance the conversation on risks at the intersection of nuclear weapons and emerging technologies.

There are many possible future directions for this conversation. Going forwards, states should consider the complex interactions between the technologies examined here. For example, the combination of technologies that 'distort' information with those that 'compress' the amount of decision-making time that human operators have available to them could have particularly disruptive effects. States should also consider how technologies will asymmetrically impact some states more than others, depending on geography, geopolitics, and nuclear status. And most important, NWS and NNWS should work together to identify how these technologies might reduce nuclear risks and seize the opportunity for cooperation when it is so desperately needed in the nuclear order.

To continue to ignore emerging technologies risks rendering the NPT irrelevant, but more importantly, to ignore emerging technologies increases the risk of nuclear weapon use. Cooperation on emerging technologies and nuclear risk reduction is necessary to reduce the likelihood and impact of a catastrophic nuclear exchange.

[144] U.S. Department of the Treasury, 'Treasury Sanctions Russia with Sweeping New Sanctions Authority', April 15, 2021, available at: https://home.treasury.gov/news/press-releases/jy0127

[145] Marion Messmer, 'Strategic Risk Reduction in the European Context', BASIC, June 2020, available at: https://basicint.org/wp-content/uploads/2020/06/Strategic-Risk-Reduction-in-the-European-Context-WEB-1.pdf

Annex A: Data collection, analysis, and validation

This Annex gives interested readers more information on the data collection, analysis, and validation processes.

Shortlisting the technologies

Whether a technology impacts nuclear risk depends on its application in the nuclear realm.[146] This study treats the application of technology as more important than any intrinsic stabilising or destabilising features of the technology itself.

This study builds off a proof-of-concept study conducted in June 2020.[147] The proof-of-concept study was also designed using the STREAM method, but due to resource limitations, it engaged a smaller expert group. The proof-of-concept study identified a longer list of technologies that are relevant to strategic stability (n=21) through semi-structured interviews and a literature review. One significant contribution of the proof-of-concept study was that it helped to identify which technologies are likely to reach an operational TRL (ie TRL 7-9) within the next 10 years.

Using the proof-of-concept study as a starting point, the list of technologies was refined from twenty-one technologies to ten. I used James Acton's typology[148] in the 'Capturing Technology, Rethinking Arms Control' volume (2020) to shortlist the 10 technologies that are most likely to impact crisis stability. Acton proposes four categories of 'nonnuclear technologies that are behind the growing danger of crisis instability', which include precise nonnuclear munitions, nonnuclear attacks on NC3, nuclear interceptors, and information-gathering capabilities. To this list, I added a fifth category – 'wormhole escalation' – which was inspired by Rebecca Hersman's article 'Wormhole Escalation in the New Nuclear Age'.[149] The addition of the fifth category expands the traditional ways of thinking about escalation, to include non-linear crisis escalation.

Finally, I held a virtual workshop in November 2020, wherein a group of subject-matter experts Red Teamed the research design. At this 90-minute workshop, experts provided their input on 1) the framing of the study, 2) the shortlist of technologies, and 3) the technology scoring criteria. This workshop enabled me to adjust, combine, and eliminate technologies from the shortlist.

Evaluating the technologies

This study utilises the Systematic Technology Reconnaissance, Evaluation and Adoption Methodology (STREAM), developed by Steven W. Popper et al.[150] As discussed, STREAM assesses emerging and established technologies according to a range of impact and implementation criteria. In the context of this study, the STREAM method culminated in a scoring exercise, wherein subject-matter experts completed a survey that asked them to evaluate 10 technologies on the impact that they might have on crisis stability, and any barriers to their implementation in the United Kingdom in the next 10 years.[151] Table 2 contains more detail on the individual questions of the technology scoring exercise.

[146] Madeline Zutt and Michal Onderco, 'How emerging technologies impact the future of nuclear risk and arms control', European Leadership Network, September 1, 2020, available at: https://www.europeanleadershipnetwork.org/commentary/how-emerging-technologies-impact-the-future-of-nuclear-risk-and-arms-control

[147] Marina Favaro, 'Cyber- and Space-Based Capabilities and Their Impact on Strategic Stability', The 2020 UK PONI Papers, October 2020, available at: https://rusi.org/sites/default/files/202011_poni_papers_2020_web.pdf#page=54

[148] James M. Acton, 'Strategic Stability and the Global Race for Technological Leadership', Capturing Technology, Rethinking Arms Control Conference Reader, November 2020, available at: https://rethinkingarmscontrol.de/wp-content/uploads/2020/10/20-AA-RAC-Reader-2020-10-28-final.pdf#page=6

[149] Rebecca Hersman, 'Wormhole Escalation in the New Nuclear Age', Texas National Security Review, 3:3 (2020), available at: http://dx.doi.org/10.26153/tsw/10220

[150] Steven W. Popper, Nidhi Kalra, Richard Silberglitt, Edmundo Molina-Perez, Youngbok Ryu, and Michael Scarpati, 'Strategic Issues Facing Transportation, Volume 3: Expediting Future Technologies for Enhancing Transportation System Performance', NCHRP Report 750, 2013, available at: https://doi.org/10.17226/22448

[151] I used existing social networks to identify the subject-matter experts, with assistance from my colleagues at King's College London. Like any non-random sampling method, 'snowball sampling' does not guarantee representation and lends itself to a community bias risk.

The guiding question underpinning the impact variables is: How will the deployment of this technology impact the stability of a theoretical crisis between the UK and another nuclear actor?[152]

On the other hand, the guiding question underpinning the feasibility of implementation variables is: What is the trajectory for the development of this technology and its implementation into UK strategic posture?

Table 2: Impact and implementation criteria

	Impact		Feasibility of implementation	
Guiding question	How will the deployment of this technology impact the stability of a theoretical crisis between the United Kingdom and another nuclear actor?		What is the trajectory for the development of this technology and its implementation into UK strategic posture?	
Scoring criteria	Q1.1	Can this technology deliver, or enable the delivery of, a disarming first strike?	Q2.1	How advanced is the development of this technology in the UK context?
	Q1.1a	If yes, to what extent could this technology be used to deliver a disarming first strike that eliminates an adversary's ability to retaliate with nuclear weapons?	Q2.2	To what extent might budgetary barriers limit the development/implementation of this technology for the United Kingdom?
	Q1.1b	If yes, to what extent does the technology confer such a significant advance in first-strike capabilities that an adversary would be more likely to launch first?	Q2.3	To what extent might human barriers (eg training requirements) limit the development/implementation of this technology for the United Kingdom?
	Q1.1c	If no, how likely is it that the use of this technology would elicit a nuclear response?	Q2.4	To what extent might regulatory/ policy barriers limit the development/ implementation of this technology for the United Kingdom?
	Q1.2	To what extent could this technology be used to strengthen or erode nuclear command, control and communications (NC3)?	Q2.5	To what extent might ethical/legal barriers (eg IHL; law of armed conflict) limit the development/implementation of this technology for the United Kingdom?
	Q1.3	To what extent could the technology increase or reduce decision-making time/ situational awareness during a crisis?	Q2.6	To what extent might technical barriers limit the development/implementation of this technology for the United Kingdom?
	Q1.4	To what extent could this technology increase mis/disinformation during a crisis?		
	Q1.5	To what extent are there credible defensive measures in place that could prevent or blunt an attack using this technology?		

[152] 'The UK's hypothetical adversary would necessarily be another nuclear actor because crisis stability is defined as a situation where no party has an incentive for nuclear first strike or pre-emption. Thus, its focus is on the size of the gap between the payoffs for striking first and striking second with nuclear weapons.' Source: Kristin Ven Bruusgaard and Jaclyn A. Kerr, 'Crisis Stability and the Impact of the Information Ecosystem', Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict, March 15, 2020, available at: https://www.hoover.org/sites/default/files/research/docs/trinkunas_threetweetstomidnight_137-158_ch.7.pdf

One of the reasons for focusing on the UK is scoping; I partitioned a small but achievable project that I could execute to a high quality. Evaluating the feasibility of implementing a given technology for 'red' (meaning adversarial) forces falls outside the scope of this study. In other words, this study does not account for the ways in which implementation drivers and barriers might be similar or different between the UK and its adversaries, nor does it account for future countermeasures. Regardless, these research findings lend themselves to various inferences. For example, one could infer that if a given technology is feasible for the UK, then it is also feasible for NATO. Alternatively, if a technology is feasible for the private sector in the UK, then one could infer that it is feasible for the private sector elsewhere.

After one month, 61 completed surveys were submitted.[153] The respondents exemplify nearly equal representation from the policy community and the S&T community (58 per cent:42 per cent), which was a primary objective of the study.[154] The average number of technologies that experts responded to was between 3 and 4.[155] This means that most participants scored a minority of technologies, which is in keeping with the instructions for the survey.[156]

Clustering the technologies

Three techniques were used to analyse the data generated by the STREAM technology scoring exercise: summation notation, Machine Learning k-means clustering, and pairwise significance testing. Each technique makes a unique contribution to the robustness of the study and strengthens our confidence in the research findings. Each data analysis technique will be explored in turn below.

Using summation notation to cluster the technologies

The most basic way to translate this data into meaningful but easy-to-digest information is to take the mean – or average – of the impact and implementation scores, using summation notation.[157] When doing this, I took care to ensure that:

- Each question is normalised;[158]
- Each question works in the same direction;[159]
- Each question is weighted equally;
- Each question is orthogonal (ie statistically independent), so that no value is double counted; and
- The set of questions comprehensively covers all aspects of crisis stability.[160]

In calculating the mean of the impact and implementation scores, Q1.1a, 1.1b and 1.1c were removed, as these are sub questions of Q1.1 and therefore not orthogonal. Q2.1 was also removed from the implementation average because Q2.1 is less about barriers to the implementation of a given technology and more about how advanced the development of this technology is in the UK context, whereas Q2.2 to Q2.6 relate specifically to the trajectory of development for this technology and its implementation into UK strategic posture.

The results are shown in Table 3 (see overleaf).

[153] Data collection occurred between 11 January 2021 and 5 February 2021.

[154] The goal was to have roughly equal expertise from both communities (however crudely defined), because of the ability for each community to complement the expertise of the other. Whereas technical experts would be better placed to answer questions on Technology Readiness Level (TRL) or evaluate the technical barriers to developing or implementing a given technology, policy experts would be better placed to assess how regulatory or legal barriers might limit the development or implementation of a given technology. To ascertain this demographic data, experts were asked to self-identify with one community or the other.

[155] Mean = 3.8, median = 3, mode = 3.

[156] In the outreach email, I emphasised that respondents should only score technologies with which they have demonstrable expertise and/or consider themselves an 'expert'. This reflects my preference that experts score a few technologies with a high degree of confidence in their responses, rather than score all 10 technologies with varying degrees of confidence. This reflects an awareness that the analysis could only be as robust as the data inputs, and that this is will always be a challenge when using a method that relies upon expert elicitation.

[157] Summation notation is a convenient and simple form of shorthand used to give a concise expression for a sum of the values of a variable.

[158] In the technology scoring exercise, the questions had different ranges of answer (eg 1 to 5, -3 to +3, etc) In the analysis, these responses were normalized to between 0 and 1 for binary questions, or 1 to 5 for Likert scale questions.

[159] In the technology scoring exercise, the questions occasionally 'act in different directions' on crisis stability (ie in some instances, a 5 means that it is likely to escalate an ongoing crisis, and in some instances a 5 means it is likely to deescalate an ongoing crisis). In some instances, the data was 'flipped' so that everything works in the same direction (ie 5 is always high impact/high feasibility of implementation).

[160] This is very difficult to guarantee in this type of analysis. It is unclear whether one could really subdivide crisis stability into completely orthogonal questions, but every effort has been taken to do so in a statistically independent manner.

Table 3: Average impact, implementation and TRL scores

	Technology	Average Impact Score	Average Implementation Score	Average TRL
1	AI-powered cyber operations	3.69	3.05	5.67
2	Hypersonic missiles	3.73	3.30	2.68
3	Rendezvous and Proximity Operations (RPO)	3.58	3.29	5.00
4	Satellite jamming and spoofing systems	3.75	3.68	5.57
5	Kinetic anti-satellite capabilities	3.84	2.68	2.33
6	Directed energy weapons	2.83	3.41	3.77
7	Swarm robotics	3.20	3.23	4.75
8	AI for Intelligence, Surveillance and Reconnaissance (ISR)	2.75	3.74	5.57
9	Smallsats for Intelligence, Surveillance and Reconnaissance (ISR)	2.44	4.29	8.20
10	Deep-fake technology	3.71	3.62	7.00

This is visualised in Figure 2, with feasibility of implementation on the x-axis, impact on the y-axis and bubble size corresponding with TRL, meaning the larger the bubble, the higher the TRL.

Figure 2: Graphing average impact, feasibility and TRL scores

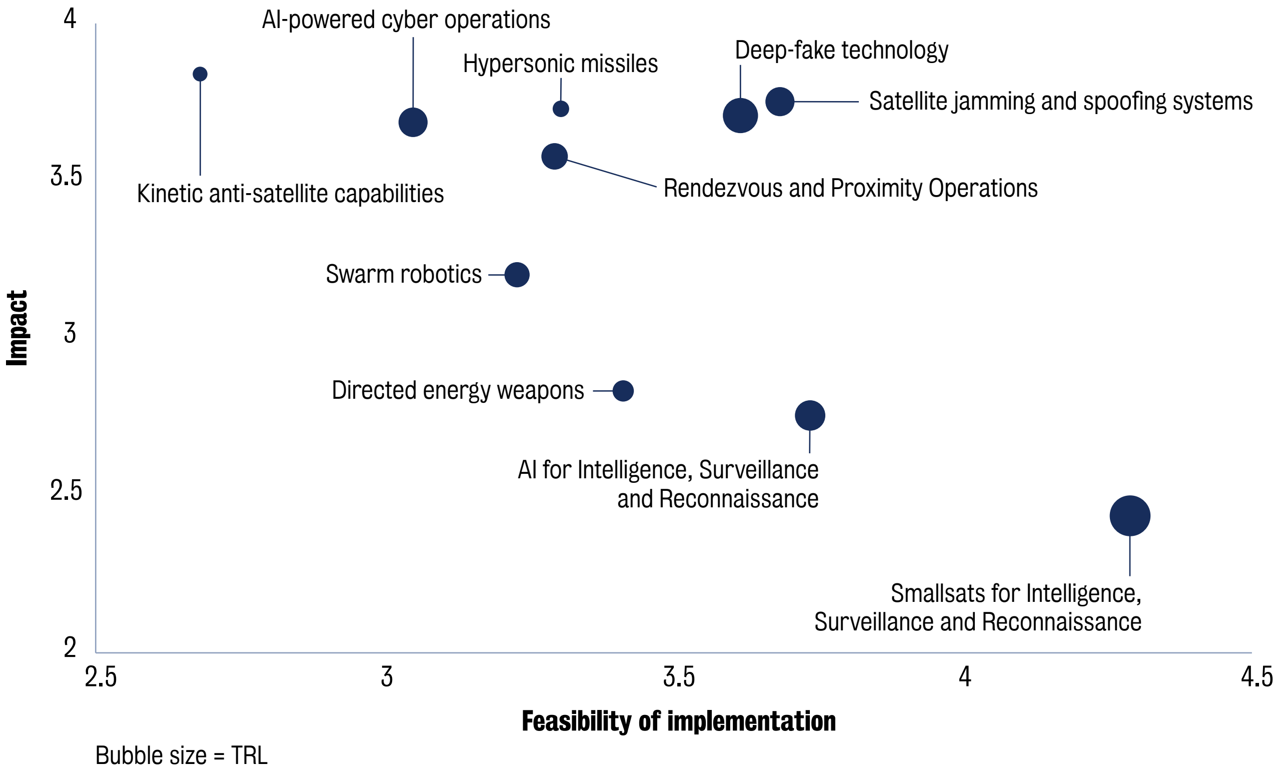


Figure 3 superimposes four quadrants, which help to make more sense of the scatter plot graph. The four quadrants show that: Cluster 1 – Distort is higher impact, higher feasibility; Cluster 2 – Compress is higher impact, lower feasibility; Cluster 3 – Thwart is lower impact, lower feasibility; and Cluster 4 – Illuminate is lower impact, higher feasibility.

Figure 3: Overlaying the four impact/implementation quadrants to get a sense of risk

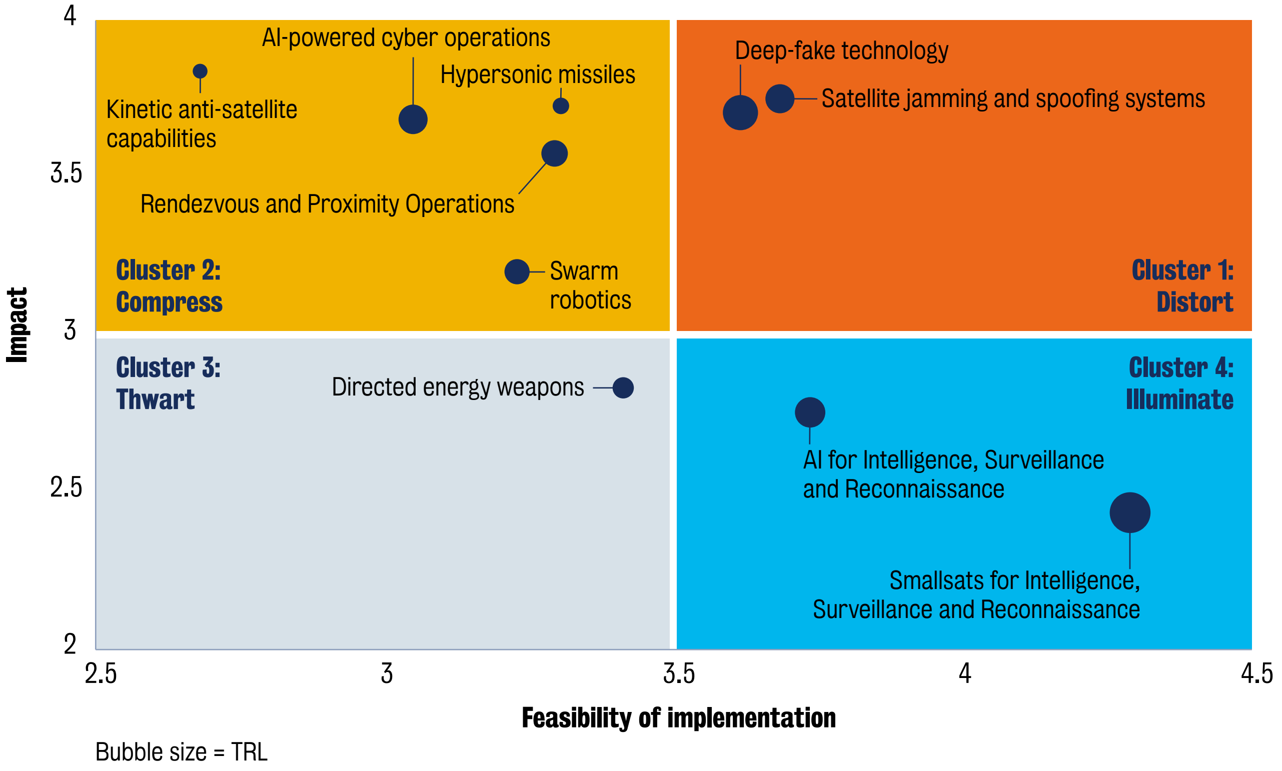
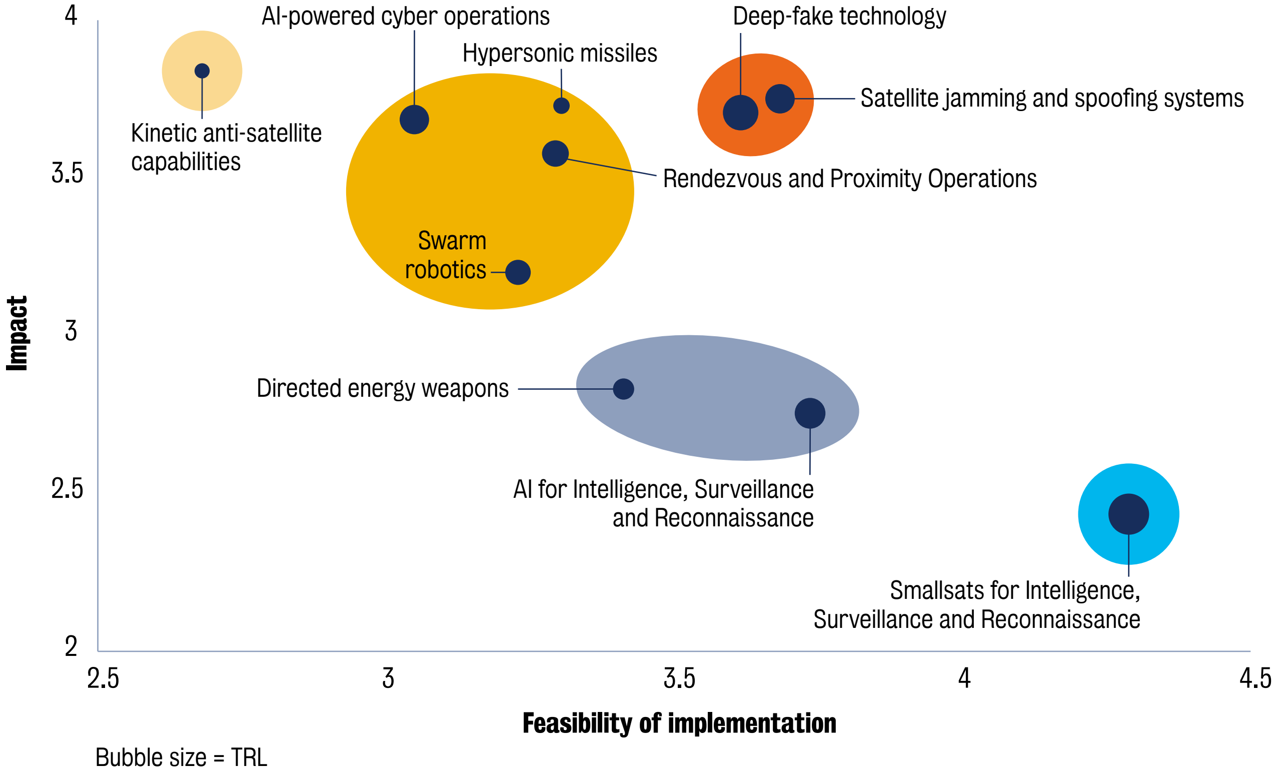


Figure 4 puts a slightly finer point on the four quadrants by attempting to 'eyeball' or approximate the technology clusters. However, this is only one of many reasonable technology clusters using summation notation.

Figure 4: Approximated technology clusters using summation notation



Using Machine Learning to cluster the technologies

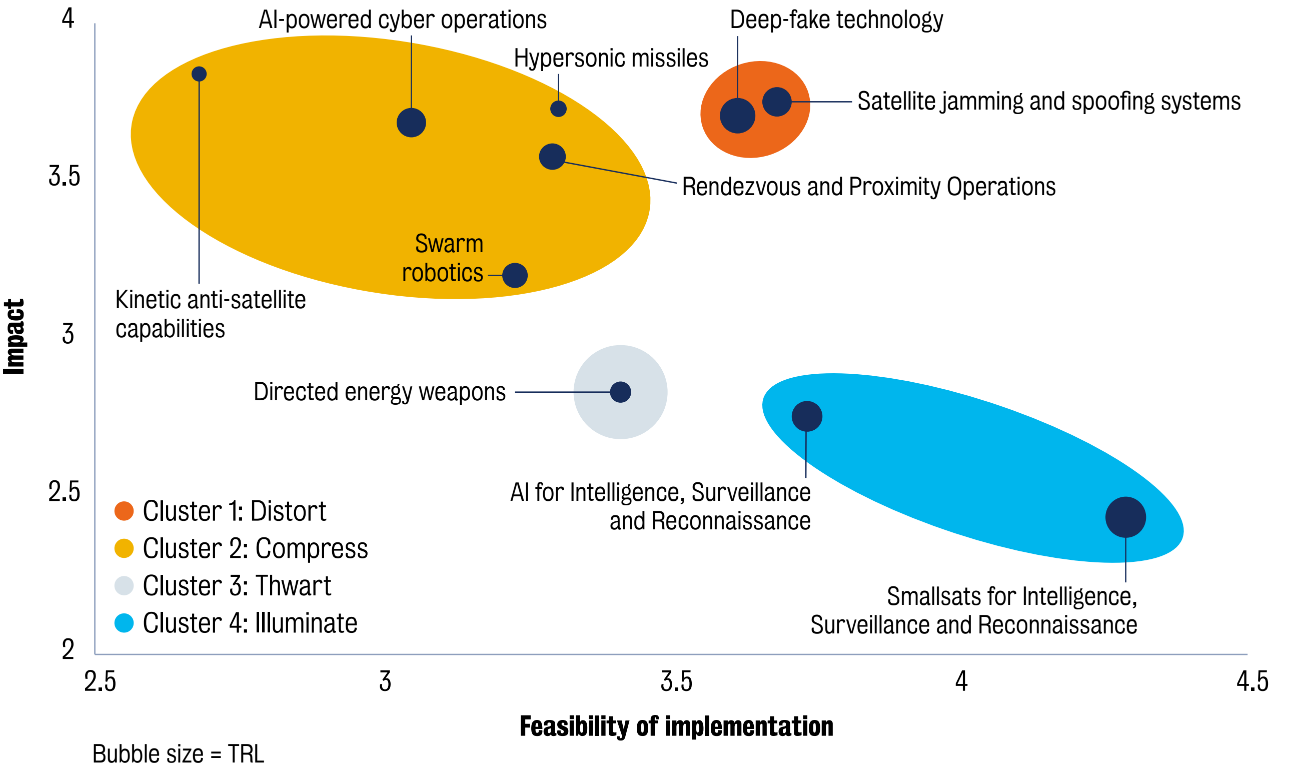
There are more robust ways of clustering technologies than summation notation. This section discusses the use of Machine Learning (ML),[161] specifically k-means clustering,[162] to this end. Rather than use the average scores, which are derived from the sum of all values of a given variable, to cluster the technologies, ML retains the value of each variable by clustering technologies on multiple axes at once.

The addition of ML does two things in the context of this study:

1. Sense-checks and validates the approximated technology clusters that were derived from summation notation; and
2. Adds more nuance. Rather than working from the mean of impact and implementation scores for each technology, the clustering algorithm groups technologies together which experts scored similarly across each of the questions, thereby preserving more nuance in the scores of individual variables. The result is that the technology clusters have a multi-dimensional depth that could not be attained by using only the average impact and implementation scores.

Although ML clusters the technologies together, the algorithm is a 'black box' to the extent that it is unknowable *on what grounds* these technologies are deemed similar or different from each other. A qualitative feedback session with a small sub-group of respondents (n=7) was used to add qualitative 'meat' – or meaning – to the computer-derived 'bones' of these technology clusters.[163] Figure 5 visualises the outcome of k-means clustering on this dataset.

Figure 5: Using ML k-means clustering to identify four technology clusters



[161] Machine Learning comprises algorithms and statistical models that computer systems use to perform a specific task effectively without using explicit instructions, relying on patterns and inference to build a mathematical model based on sample data.

[162] In k-means clustering, 'k' refers to the number of clusters. When selecting the number of clusters, the objective is to minimise variation in the clusters up to the point of diminishing returns to scale.

[163] The qualitative debrief was held virtually via Zoom on 4 March 2021.

These four technology clusters form the basis of the analysis in the main body of the policy paper, because they are the most robust.

Using statistical significance testing to validate the technology clusters

Finally, pairwise statistical significance testing was used to validate the research findings. This technique determines whether the scores for each technology are sufficiently different from each other. In other words, statistical significance testing confirms that the difference in scores comes from actual differences in the characteristics of the technologies, rather than random chance.

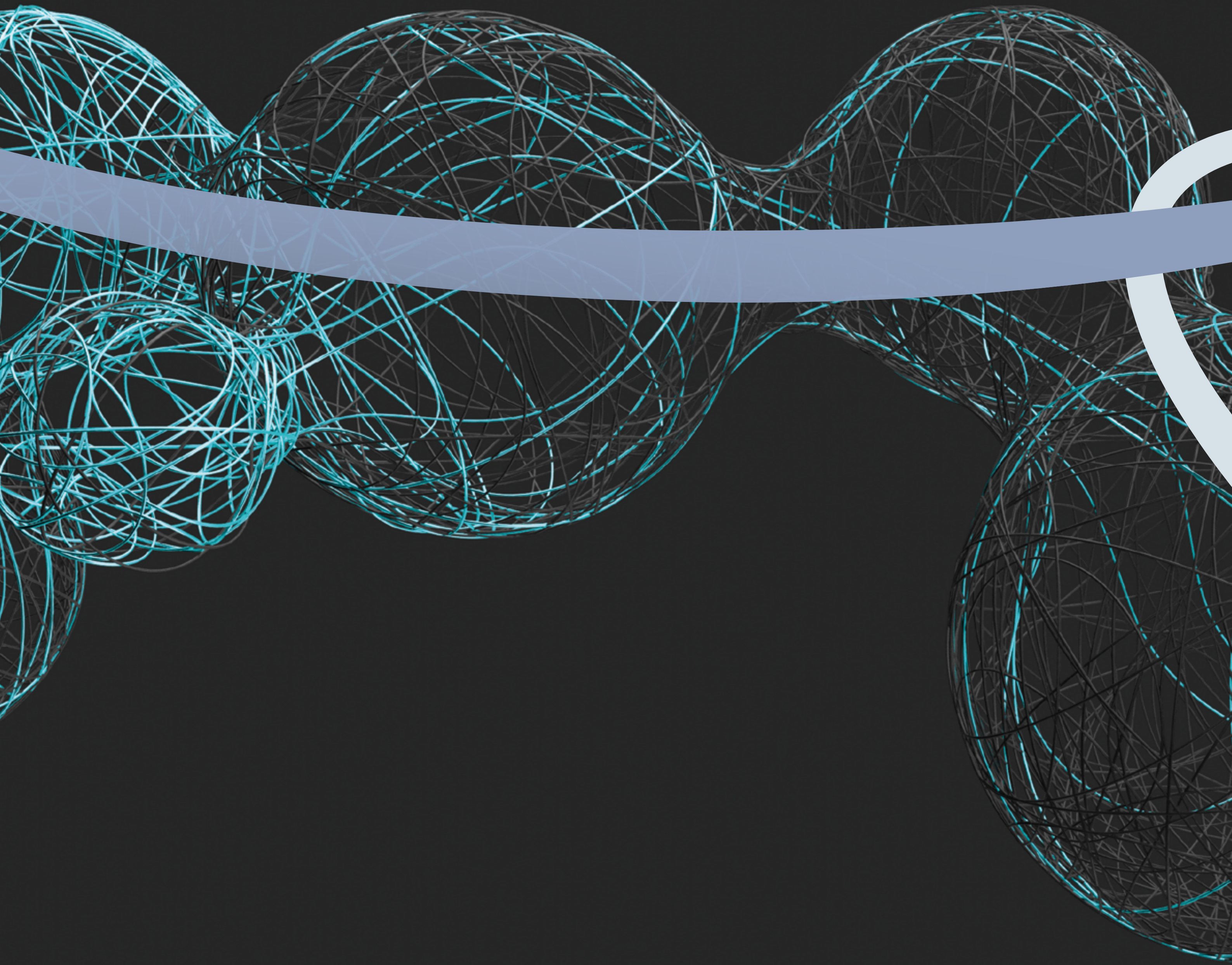
In each technology cluster, some of the technologies were scored differently and some were scored the same. However, for technologies in different clusters, the differences in scores were always found to be statistically significant.[164] For example, if Technology A is in Cluster A and Technology B is in Cluster B, then their scores are always statistically different. This further demonstrates that the clusters represent distinct and recognisable groups of technologies. It also highlights differences between technologies in the same group, which is to be expected. These findings do not suggest that all the technologies in each cluster are identical, merely that they have more in common with each other than with other shortlisted technologies. Finally, the combined scores for each cluster were compared to each other using pairwise t-testing.[165] The combined scores of technology clusters were always found to be statistically different. This strengthens our confidence in the technology clusters.

[164] The level of significance was set at 0.05, but in reality, all technology pairs (where the technologies are in different clusters), had a significance level of less than 0.01, except in the case of Rendezvous and Proximity Operation and deep fake technology. The significance level for this pair is 0.018.

[165] The Likert scale is ordinal (ie 1 – To a very low degree to 5 – To a very high degree) and the t-test was initially developed to use with quantitative variables that have a normal distribution. However, scholars have made the case that it is statistically acceptable to test the difference of means using a t-test when the variable is a Likert scale, and the population does not have a normal distribution. Source: Pedro Cosme Costa Vieira, 'T-test with Likert scale variables', SSRN, April 26, 2016, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2770035

Acknowledgements

My sincere thanks to Dr Heather Williams, who provided me with the space, resources, encouragement, and sounding board that I needed to complete this research. For their insightful comments on previous drafts of this article, I would like to thank James Black, Alexandra Stickings, and Douglas Shaw. I am deeply grateful to Richard Flint for his data science and data visualisation contributions. I would also like to acknowledge my colleagues at King's College London – Jamie Kwong, Erin Connolly, and Amelia Morgan – for their support and advice at key points in this project. Finally, thanks to everyone who volunteered their time and expertise to this study as an interviewee, Red Team workshop participant, technology scorer, or qualitative exercise debrief contributor. This paper contributes to an ongoing project at King's College London with the generous support of the MacArthur Foundation.



Centre for Science and Security Studies
Department of War Studies
King's College London
Strand
London WC2R 2LS
United Kingdom

www.kcl.ac.uk/csss
@KCL_CSSS

Designed by: RF Design (May 2021)
Approved by: brand@kcl.ac.uk