

CSSPR: Center for Security, Strategy
and Policy Research

Looking Beyond the Risk:

The Advantages of “Unavoidable” Technologies



NUCLEAR FELLOWSHIP
PUBLICATION



THE
UNIVERSITY OF
LAHORE

Author

Angelo Frasca Caccia is a CSSPR Nuclear Scholars Fellow of the 2021 cohort. Angelo holds a Bachelor's degree in Investigation and Security Sciences from Università degli Studi di Perugia and recently graduated in International Security Studies with distinction from the Department of Politics and International Relations (HyPIR) of the University of Leicester. His research revolves around escalation dynamics and nuclear deterrence between nuclear powers, as well as around the evolution of the traditional Cold War concepts of escalation ladder and escalation dominance.

© The Center for Security, Strategy and Policy Research 2021

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of CSSPR or UOL.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical including photocopying, recording or any information storage or retrieval system, without the prior written permission of the copyright holder.

Please direct all enquiries to the publishers. All imagery in this report is licensed under Creative Commons 2.0 from Unsplash.com, with thanks to all the photographers.

CSSPR

The Centre for Security, Strategy and Policy Research (CSSPR) is an autonomous policy research center housed in the University of Lahore (UOL), Lahore. It aims at generating research on issues relating to contemporary security and strategic aspects that affect Pakistan's domestic security environment and its international relations with other countries of the world. The Center aims to generate research-driven and nuanced discourses on how best to address security and strategic challenges that deny Pakistan the opportunity and ability to play its role as a responsible seventh Nuclear Weapon State in the world.

The Center conducts research, publishes, and holds conferences on a broad range of regional and international security, strategic, peace and conflict issues. Through this Center, we aim to establish national, regional and international linkages and collaborate with universities, think tanks, research institutes and eminent scholars from around the world, enabling national and international scholars and our students to use CSSPR's platform to deliberate upon issues that carry a global impact.

Center for Security, Strategy and Policy Research (CSSPR)

1-Km Defence Road
Lahore
Email: csspr@siss.uol.edu.pk
www.csspr.uol.edu.pk

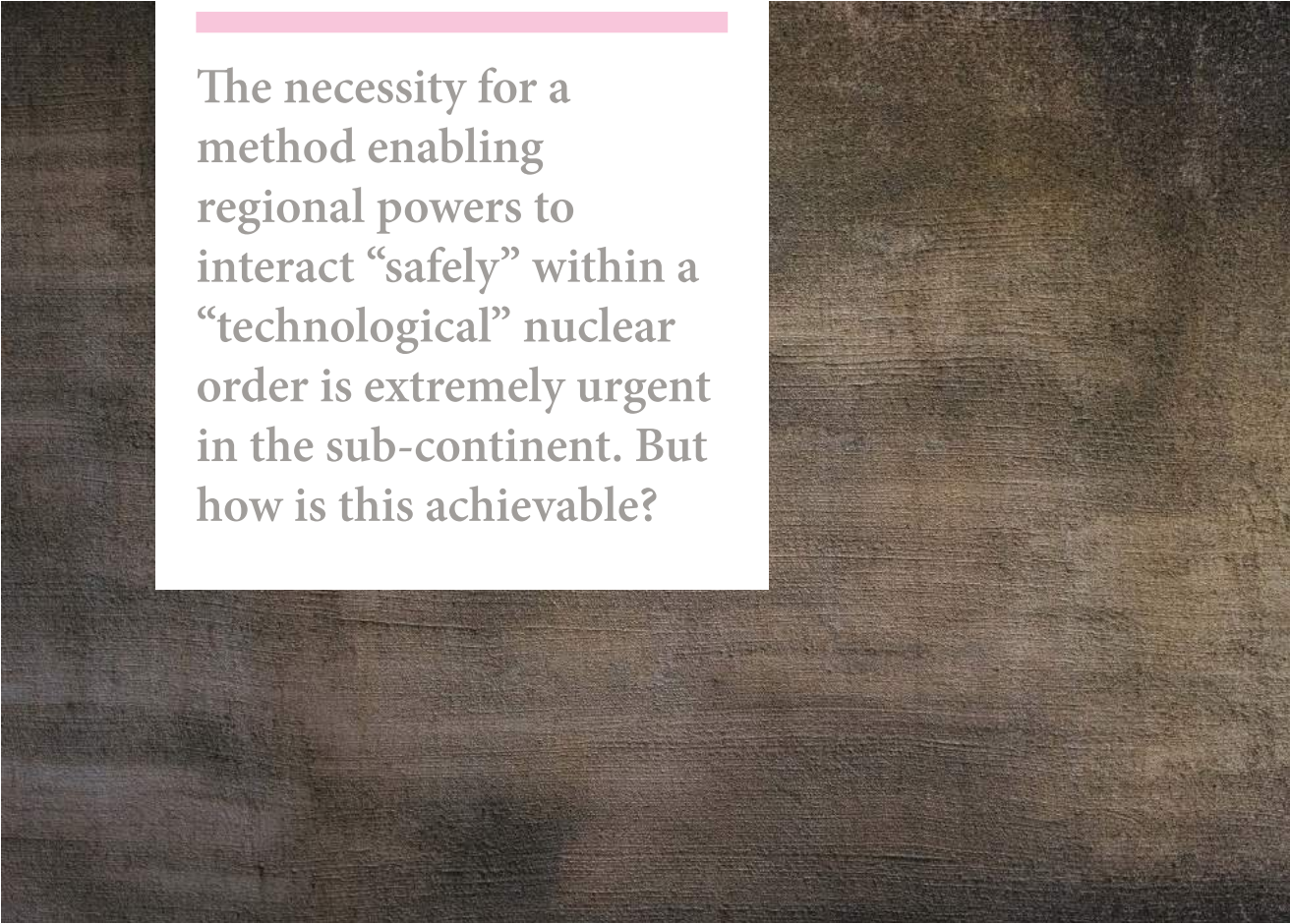
University of Lahore (UOL)

1-Km Defence Road
Lahore
Email: info@uol.edu.pk
T: +92 423 2233888
www.uol.edu.pk

Looking Beyond the Risk: The Advantages of “Unavoidable” Technologies

There is widespread acknowledgment that the increasing reliance on non-nuclear technologies with strategic applications exerts a destabilizing effect on the global nuclear order, especially if one thinks that the implementation of the rising Artificial Intelligence (AI) can potentially provide these developments with a frightening degree of autonomy. From this perspective, even sub-conventional crises could suddenly escalate to high levels of intensity through unexpected and unintended escalation dynamics, where the failure of the rational practice of deterrence gets entangled with the complexity of non-human, artificial decision-making. Of course, this is a single, worst-case scenario lacking a “human in the loop”, but contemporary deterrence strategies are arguably weak in the face of the emerging, asymmetric escalation landscape. According to Rebecca Hersman, “holes may suddenly open in the fabric of deterrence”, where low-intensity crises can escalate to the strategic level through so-called “wormhole” escalation dynamics. One example is disinformation campaigns, which could prompt a leader to take action prematurely and inadvertently escalate a dispute. Unintended escalation is therefore going to become more relevant as global actors increasingly rely on strategic non-nuclear weapons (SNNWs) to exercise cross-domain coercion. What is more, such a growing reliance on SNNWs may even indicate that a new Nuclear Age is approaching fast, as these technologies are causing a paradigm shift from a global nuclear order where nuclear weapons dictated the terms of stability and instability, to one in which this role is mostly played by SNNWs.

Equally, these sorts of issues have a considerable impact on the South Asian security environment as well. In particular, regional escalation dynamics between Pakistan and India often take the form of perilous brinkmanship where neither side is willing to concede to the other, thus pushing the dispute to the very edge and forcing the United States to intervene, dampen tensions and de-escalate the crisis. As such, this type of de-escalation trend became operational when the risk of nuclear escalation was evidently high, especially following a gradual exacerbation of tensions and Pakistan signalling nuclear danger. However, the looming threat of inadvertent, wormhole escalation dynamics causing strategic intensification of sub-conventional crises clearly implies that India-Pakistan brinkmanship could end up with



The necessity for a method enabling regional powers to interact “safely” within a “technological” nuclear order is extremely urgent in the sub-continent. But how is this achievable?

a nuclear exchange before de-escalation by foreign intervention comes into play, because even low-intensity crises could now escalate to unbearable levels in an unthinkable and unpredictable manner. For this reason, the necessity for a method enabling regional powers to interact “safely” within a “technological” nuclear order is extremely urgent in the sub-continent. But how is this achievable?

Technological Opportunities

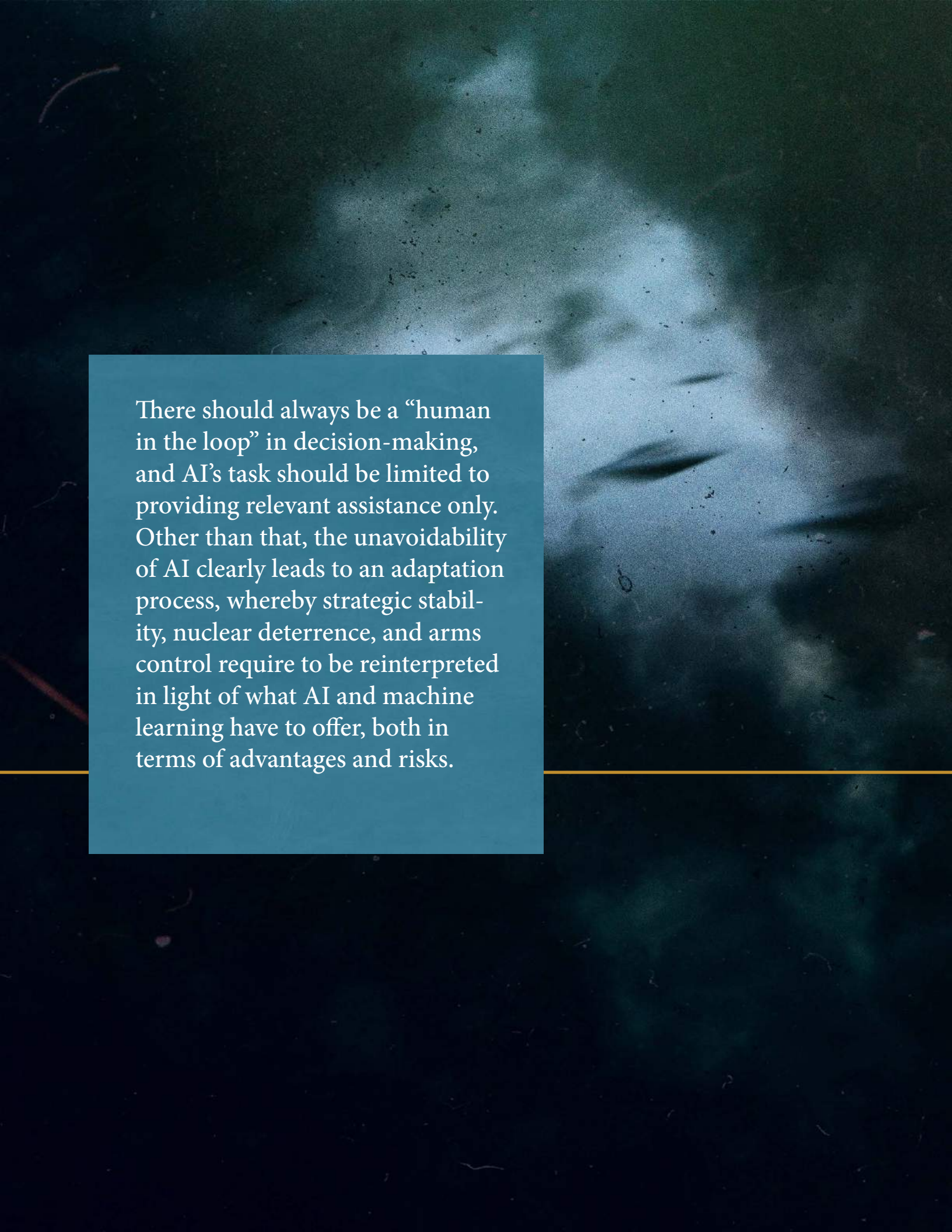
Countering the proliferation of emerging technologies is definitely not an option, as technology enablers like Artificial Intelligence are often referred to as indispensable and unavoidable as electricity. On the other hand, a social constructivist perspective would suggest that if the proliferation of non-nuclear technologies with strategic applications stems from an impulse to deliberately escalate crises, then a willingness to de-escalate them may perhaps follow the same procedure. In other words, because the specific degree of risk related to a technology depends on its application, we may wonder whether such technology could be appositely and smartly exploited to actually maintain or possibly enhance strategic stability, or even de-escalate crises. Relatedly, the air battle following the 2019 Balakot strike represents an interesting case-study, as India claimed that a MiG-21 fighter jet actually destroyed a Pakistani F-16 before it was downed and its pilot captured,

even though evidence demonstrated that none of Pakistan's F-16s was missing. According to Rebecca Hersman such an "auto-declared victory" may have borrowed time in terms of stability, and consequently questions would be raised on whether states can "secure the benefits of de-escalation while forgoing the political costs of military defeat by creating and propagating their own alternative, and even incompatible, perceptions of victory". In such a scenario, disinformation campaigns and the spread of tailored fake news can certainly play a role, but according to Hersman, this whole approach looks more like "a house of cards ready to collapse even more precipitously with the next crisis".

In contrast to the targeted spread of information, technology enablers may be more suitable for a de-escalation or strategic stabilizer role. In a very recent article, Jessica Cox and Heather Williams delineate at least two fields where the implementation of AI can provide a positive effect. First, nuclear deterrence and strategic stability can be improved by integrating AI in early-warning systems and decision-making analysis, with potentially positive implications for information-processing mechanisms and de-escalation timings in crisis situations. As the authors note, in 1995 the radar crew of the Russian Missile Attack Warning System detected a Norwegian civilian research rocket launch and mistook it for a US Trident II SLBM. The disaster was avoided because Russian authorities eventually concluded that this was not a ballistic missile after carefully reviewing satellite information. Relatedly, the implementation of AI in early-warning systems can help mitigate these eventualities through provision of a better quality of information and more accurate analyses, thus assisting human decision-making when the risk of escalation is significant. Second, "smart exploitation" of AI may potentially benefit arms control regimes by improving verification mechanisms. For example, a possible implementation field for AI is that of missile tracking, which could provide additional verification tools in arms control agreements.

Benefits for South Asia

Clearly, the solutions aforementioned can be adopted in the South Asian security environment as well. The area of confidence-building measures (CBMs) is certainly an important portion of India-Pakistan relationship, where a number of agreements are currently in place to deal with the risk of escalation in crisis situations. For example, a 1988 agreement prohibited attacks against nuclear facilities on both sides through provision of lists of respective nuclear facilities. Similarly, in 2005, a missile test pre-notification agreement was signed, and in 2007 the risk of accidents and unauthorized nuclear use became the object of another agreement. However, it should be also noted that none of these bargains is actually ruled by a proper verification mechanism. For this reason, Artificial Intelligence could arguably provide a better and more accurate insight about the locations of both countries' nuclear facilities through ameliorating satellite imagery for instance. This would also help both sides mitigate the sensitivities regarding in-person intrusive monitoring, which is one of the primary obstacles when it comes to the establishment of verification systems and procedures. Similarly, South Asian security environment could considerably benefit from the advantages that AI has to offer in terms



There should always be a “human in the loop” in decision-making, and AI’s task should be limited to providing relevant assistance only. Other than that, the unavoidability of AI clearly leads to an adaptation process, whereby strategic stability, nuclear deterrence, and arms control require to be reinterpreted in light of what AI and machine learning have to offer, both in terms of advantages and risks.

of strategic stability and nuclear deterrence. In particular, it is widely acknowledged that the flight time for missiles would be just a few minutes in case of an exchange between the two regional powers. Consequently, a classic crisis situation would plausibly entail short decision-making timelines, which are also likely to translate into either intentional or inadvertent nuclear escalation if coupled with the “fog of war”. AI can accordingly provide a better insight on whether a launched missile is delivering a conventional or nuclear payload through a quick assessment of telemetry and technical information for instance given that the majority of both India’s and Pakistan’s missiles are dual-capable. Alternatively, detected launches can be assessed against the regular monitoring of activities, thus helping decision-makers identify anomalous attacks and quickly adjust their response in cases of immediate threat. This could also be an option when identifying launches not complying with the 2005 Missile Notification Pact. In other words, the promptness of AI-based calculations could assist decision-makers when the imminence of a threat, also coupled with a certain degree of ambiguity, does not allow time for verification, targeted response, and de-escalation.

Conclusion: Risks and Optimization

As nuclear powers seek to make the most of AI-related benefits in arms control regimes and decision-making systems, such a process should be also met and optimized by an understanding of the risks associated with AI and how to mitigate them. In general, any kind of technology lacks human sensibility, and predictions of signalling, for instance, would be related to recognition of recurring behavioural patterns. On the other hand, if new dynamics play out or the behavioural pattern suddenly occurs with specific exceptions, then predictions would be more difficult for AI-based systems, and human intervention would be required to avoid inadvertent escalation. For example, India-Pakistan brinkmanship is often characterized by India emphasizing conventional readiness while Pakistan accurately manipulating nuclear danger to call back US attention and start de-escalation. Accordingly, a plausible response by India’s AI-led decision-making systems could interpret Pakistan nuclear signalling as actual willingness to use nuclear weapons, rather than viewing that behavioural pattern as a specific, exceptional procedure that aims to encourage de-escalation by foreign intervention. For this reason, there should always be a “human in the loop” in decision-making, and AI’s task should be limited to providing relevant assistance only. Other than that, the unavoidability of AI clearly leads to an adaptation process, whereby strategic stability, nuclear deterrence, and arms control require to be reinterpreted in light of what AI and machine learning have to offer, both in terms of advantages and risks.

References

Cox J. & Williams H. (2021) The Unavoidable Technology: How Artificial Intelligence Can Strengthen Nuclear Stability, *The Washington Quarterly*, 44:1, pp. 69-85;

Futter A. & Zala B. (2021) Strategic non-nuclear weapons and the onset of a Third Nuclear Age, *European Journal of International Security*, pp. 1-21;

Hersman R. (2020) Wormhole Escalation in the New Nuclear Age, *Texas National Security Review*, 3:3, pp. 91-109;

Jaffery S. A. Z. (2019) Sav review: The people behind the guns: the relationship between cutting-edge technology and escalation. Available at: <https://southasianvoices.org/the-people-behind-the-guns-the-relationship-between-cutting-edge-technology-and-escalation/> (accessed: 09/05/2021);

Khan F. (2003) The Independence-Dependence Paradox: Stability Dilemmas in South. Available at: <https://www.armscontrol.org/act/2003-10/features/independence-dependence-paradox-stability-dilemmas-south> (accessed: 09/05/2021);

Krepon M. (2017) South Asia Confidence-Building Measures (CBM) Timeline. Available at: <https://www.stimson.org/2017/south-asia-confidence-building-measures-cbm-timeline/> (accessed: 10/05/2021);

Krepon M. and Dowling L. (2018) “Crisis intensity and nuclear Signaling in South Asia”, in Lalwani S. (ed.) et al, *Investigating Crises: South Asia’s Lessons, Evolving Dynamics, and Trajectories*. Stimson: Washington DC, pp. 187-218;

Kristensen H. M. & Korda M. (2018) Indian nuclear forces, 2018, *Bulletin of the Atomic Scientists*, 74:6, pp. 361-366;

Kristensen H. M. et al (2018) Pakistani nuclear forces, 2018, *Bulletin of the Atomic Scientists*, 74:5, pp. 348-358;

Robinson Snowden M. (2019), Probabilistic Verification: A New Concept for Verifying the Denuclearization of North Korea, *Arms Control Today*, 49:7, pp. 18-23;

Topychkanov P. (2020), “Artificial intelligence and strategic stability in South Asia: New horses for an old wagon?”, in Topychkanov P. (ed) et al, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*. SIPRI: Sweden, pp. 39-45.