Future Warfare: AI, Deterrence and Escalation

By
Shaza Arif

Published: November 19, 2026

The interaction of artificial intelligence (AI) and nuclear deterrence continues to shape the discourse surrounding future warfare. Historically, nuclear deterrence has relied on the interpretation of adversary's intent, credible retaliation, and room for human judgement. In contrast, AI alters these conditions by rapid data processing and algorithmic assessment of impending threats. Hence, in the interplay of the two central aspects of modern warfare, the main issue remains whether their convergence and the subsequent decision making enhance or weaken the conditions for strategic stability.

In traditional warfare, nuclear deterrence is highly influenced by perceptions of what military abilities can achieve. AI-enabled capabilities, when viewed in the context of the nuclear operations, weaken confidence in the survivability of mobile forces. When there is a widespread perception that second-strike assets are more vulnerable, the traditional logic of mutual vulnerability is destabilized. Decision makers may inherit a pre-structured perception competing with logic of restraint. In such a context, the pre-emption dilemma becomes structural. Resultantly, the advantage of striking first amplify, even when neither side wants to escalate further. In addition, algorithmic feedback can add new layers of complexity via mutually reinforcing threat assessments. The cumulative impact of such a scenario is the rapid intensification of the security dilemma and the operationalization of the stability-instability paradox.

On top of this, as seen in the recent Starlink episode in the Russia-Ukraine conflict, external service providers can unilaterally withdraw access to AI infrastructure. Hence, excessive dependence on external services may leave states exposed to vulnerabilities at critical moments, further complicating deterrence dynamics. Hence, the same technologies intended to enhance defensive awareness might generate catastrophic results on the battlefield. The danger is compounded by the fact that even technically accurate observations can lead to incorrect conclusions.

When the interplay of AI and nuclear deterrence are discussed, the challenge is not merely speed alone. It is the epistemic authority and perceptions that algorithmic outputs amplify. The ongoing war in the Middle East reflects how AI-assisted targeting in conventional conflicts can operationalize these concerns due to misread patterns and incorrect perceptions. The attack on the school in Minab stands as a prominent example regarding the reliability of such systems in time-sensitive windows. If such inputs are treated as authoritative, the space for independent judgment contracts significantly, unfolding catastrophic consequences. The black box nature of AI systems further intensifies threats. Hence, algorithmic recommendations should never become the unquestioned foundation for offense in sensitive situations as it limits the range of politically acceptable responses. The broader significance lies in the distinction between ensuring the contextual validity of AI-enabled decisions. In nuclear environments, such interpretation is consequential as any incorrect assessment can shape a response before the underlying assumption can be scrutinized. This development becomes particularly acute in states and regions with geographic proximity and hostile undercurrents such as in South Asia (India - Pakistan) and Eastern Europe (NATO-Russia).

In complex, volatile, and technologically enabled environments, retaining meaningful human judgement is pivotal to mitigate the present dangers. To be fair, ensuring meaningful human control requires more than placing a human as decisionmaker at the final stage of a process. During compressed timelines, the formal authority may coexist with diminished control. Hence, meaningful human control requires the distinction between authorization control and cognitive control. The former refers to the formal authority to approve an action, whereas the latter refers to the conditions that enable independent assessment, interrogation of the assumptions, and analysis of the present options and rejection of any escalatory command. In practice, the larger risk is that authorization control formally remains intact, while cognitive control gradually deteriorates under the pressure of machine-generated urgency. The principal challenge is not the outright autonomous launch of nuclear weapons, but the normalization of machine-speed decision making and subordination of independent judgment, opening the way for inadvertent escalation.

In such a complex and clouded environment, preserving strategic stability requires evaluating the cumulative impact on deterrence rather than measuring its value exclusively via speed, precision, and operational efficiency. Future conflicts will increasingly involve interconnected operations across conventional sources, cyber networks, outer space, and information infrastructure. In such a context, the doctrinal distinction between nuclear and non-nuclear functions becomes more complex. Crisis communication channels, both technical and diplomatic, must be designed to operate under conditions of partial system degradation. Without such channels, reliance on algorithmically confident assessments can increase. Furthermore, this approach must be complemented by robust verification protocols, multi-source intelligence, independent analytical pathways, and an institutional refusal to treat algorithmic recommendations as cognitive authority. Lastly, human control should be treated as central to a procedural requirement. These safeguards should be seen as essential components in volatile and sensitive environments.

The future of warfare will be influenced by the decision-making infrastructures. The ultimate challenge is to reconcile technological speed with cognitive control. In nuclear operations, the primary advantage lies in maintaining the capacity to remain deliberative when every signal points towards urgency.

Shaza Arif is a Senior Research Associate at the Centre for Aerospace & Security Studies, (CASS), Islamabad. She can be reached at cass.thinkers@casstt.com. The views of the author are her own.