ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

# AI AND PREEMPTIVE STRIKE LOGIC: DOES FASTER DECISION-MAKING INCREASE THE LIKELIHOOD OF MISCALCULATION

***Usama Rasheed***
*M.Phil Scholar, Department of International Relations, University of Okara*
*Email: chusamarasheed096@gmail.com*
***Dr. Fakhara Shahid (Corresponding Author)***
*Assistant Professor, Department of International Relations, University of Okara*
*Email: fakhara.shahid@uo.edu.pk*

***Abstract:***
*The introduction of AI in the military command and control has fundamentally reversed the nature of warfare in contemporary contexts, particularly in preemptive attacks. I am examining in my research whether the use of AI-powered decision-making indeed increases the risk of strategic errors during times of crisis. I would be using a qualitative, contextual approach, digging in to the historical examples of such situations as the Cuban Missile Crisis and Cold War false alarms, as well as the present applications of AI in the military, such as autonomous targeting, early warning, and data fusion. I even created a fake AI-versus-AI crisis case study to follow the escalation process, how the threats are perceived in a high-speed atmosphere, and how the algorithms may respond in the situation when the decision is made in a super-fast manner. The findings demonstrate that although AI will be able to enhance situational awareness and reduce response times, its rapidity, when it is not monitored by a human, has the side effect of causing the perception of inaccurate signal interpretation, the inadvertent increase in the tense state of affairs, and the massacre of opportunities of intelligent diplomacy. Another red flagged issue of the research is that increasing the pace of things may even spoil strategic stability. What I propose is close AI validation, multibillion human control and novel technological (or theological, the original expression is quite bizarre) to manage the explosion risk. The contribution of this work to the existing body of knowledge about AI-based crisis management is that it is an area of interest to defense planners, policymakers, and scholars who struggle to strike a balance between the benefits and drawbacks of technologies and their impact on strategic stability.*
***Keywords:*** *Artificial Intelligence, Preemptive Strike, Crisis Stability, Strategic Miscalculation, Autonomous Weapons, Early Warning System.*

## 1. Introduction

Technological innovation, strategy doctrine, and human judgment have always been the key factors in the development of warfare. Since the industrialization of armies in the nineteenth century through the nuclear revolution in the twentieth century, improvements in the capability of military forces have not only changed the destructive ability of the state but the speed and nature of decision making during wartime. Every technological transformation has changed the connection between information perception and action. The adoption of Artificial Intelligence (AI) in military command and control (C2) systems is one of the most important changes in this historical developmental trend in the modern world. Contrary to the past where innovation was mainly used to improve firepower or maneuvering, AI directly interferes in the cognitive aspect of warfare by increasing the speed at which data is processed, threat evaluated, and functioning suggested. This change has far reaching implications on the issue of crisis stability, deterrence and preemptive logic of strikes.

The doctrine of preemptive strike, or, in other words, the use of force to avert a threat that is likely to happen before it occurs, has traditionally taken center stage in the strategic literature and discussions on security (Mueller, K. P., Castillo, J. J., Morgan, F. E., Pegahi, N., & Rosen, B. 2006). The justification of preemption is basing on the concept of vulnerability: in the case, when a state feels that it is vulnerable because of waiting, it might decide to do the first strike

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

before the state would gain or sustain its advantageous situation. Traditionally, the complicated process that entailed intelligence interpretation, inter-agency discussion, and political sanction has been adopted to make such decisions. Human decision makers have used contextual judgment, strategic signaling and diplomatic back channels to deal with escalation even in the face of extreme pressure. Nonetheless, with the implementation of AI in C2 systems, the process is radically redefined by condensing the Observe-Orient-Decide-Act (OODA) loop to allow the generation of threats and response nearly instantly (Chang, B. A. 2021). Although the speed of processing could enhance reaction to the real threats, it also reduces the possibility to reflect and creates the risk of misinterpretation.

The crises of the past give some warning signs on the risks of making fast decisions during periods of uncertainty. In the case of the Cuban Missile Crisis in 1962, misunderstandings, incomplete information and military tensions led to the two superpowers the United States and Soviet Union being at the point of nuclear war (Kokoshin, A. A. 2007). It was political restraint, deliberation, and communication that ultimately defused the crisis and subdued the situation, as opposed to automated systems, rapid retaliation, and the like. Equally, the Cold War instances of false-alarm, most infamously the Soviet early-warning failure of 1983, showed how human operators could have prevented the incident, otherwise triggered by technical errors and the automated warning system, which would have resulted in the retaliation of unprecedented proportions. These examples highlight an important fact that human hesitation, skepticism, and contextual awareness are stabilizing factors in a high-risk setting. Conversely, AI based systems that are running at machine speed cannot be aware of the larger political and strategic context that can perceive ambiguous signals with due caution.

The modern AI applications in the military are not merely just automation. Unmanned targeting systems are able to recognize and attack threats with little human involvement; predictive analytics systems can evaluate patterns of adversarial behavior; and built-in early-warning networks can synthesize satellite, radar installations, and cyber-monitoring facilities into inference of a threat. These features contribute to the situational awareness and the effectiveness of operations to a large degree. Algorithms such as data-fusion are able to combine two or more streams of intelligence in real time providing commanders with a detailed picture of the battlefield, which would have been inaccessible in previous periods. However, the same systems are prone to technical breakdowns, biased data, spoofed signals, and cyber-attacks. In crisis environment with high speeds, even small mistakes can spread quickly and cause feedback loop of growth before the damage can be fixed.

Compression of the decision cycles is a strategic contradiction. On the one hand, deterrence can be enhanced because quicker decision-making can be taken to indicate readiness and minimize vulnerabilities to surprise attacks. Conversely, more rapid action can increase the risk of instability by compelling the taking of premature action, using incomplete or misinterpreted information. Deterrence theory lays an emphasis on credibility and communication; in case AI systems react to the perceived threats but diplomatic signaling fails to make the intent known, the possibility of unintended escalation becomes higher (Zieliński, T. 2025). In addition, there is the stability-instability paradox which holds that strategic stability at the nuclear level may exist along with volatility at low levels of conflict. This volatility can be compounded in AI-mediated settings where the automated systems work in fast processes of action and response and may override historically established restraints.

The other important issue is the bias of automation and reliance on the results of the algorithms. Human-machine interaction studies have shown that operators can trust automated systems too much, especially in cases where the automated system has a track record of technical accuracy (De Visser, E. J., Pak, R., & Shaw, T. H. 2018). This bias might diminish the distrust individuals may have in AI-generated threat evaluations even when the actual threat evaluation

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

is crafted on incorrect information in the military. In addition, AI systems that are trained on past data can carry along some underlying assumptions which cannot keep up with new strategic behavior. The enemies can also intentionally take advantage of the algorithmic weaknesses by deceiving or manipulating the computers and this poses additional risks of inaccurate calculations.

These dynamics above then bring about a basic strategic predicament of convergence between AI and preemptive strike logic. This technological acceleration that contributes to the effectiveness of operations can also result in the loss of the deliberative space that is needed to manage the crisis. The key question that shall guide the current study thus comes out with a sense of urgency; does faster AI-assisted decision-making within the framework of preemptive strikes increase the probability of strategic miscalculation during military crises? This question is at the place of intertwining of technology, strategy and policy. It involves looking beyond the technical attribute of AI systems and also the interplay between them and deterrence incentives, threat perception and escalation dynamics.

To answer this question, a number of associated dimensions will be considered. The first question is: in what ways do AI-powered systems transform the perception of threats and risk evaluation in contrast to the procedures managed by humans? Second, does faster algorithms increase deterrence by decreasing uncertainty or does it increase instability by promoting benefits of first-mover advantages? Third, what are the safeguards, both doctrinal, technical and institutional, which might be used to reduce the risks involved in high-speed automated decision-making? All these sub-questions add up and lead to the realization of whether AI is a stabilizing or destabilizing factor in a crisis setting.

This paper aims to connect past experience of human-controlled crises with new realities of AI-controlled warfare. Through a combination of the knowledge of strategic theory, the historical analysis of cases, and modeling scenarios, the research will assess the interaction between the speed of algorithms and human control and the vulnerability of the system. It is not aimed at disregarding the development of technology, but on evaluating its impact on strategic stability in a strict and analytic way. With the introduction of AI systems into the National security systems, policymakers, defense planners, and the scholars should understand how the systems may influence the preemptive strike decisions.

## 2: Research Objectives

- To investigate cases of crises in the past (e.g., close nuclear accidents) to identify the impacts of the speed of decisions and human restraint on strategic outcomes.
- To test the implementation of AI in contemporary command and control in the military, specifically autonomous targeting, early warning, and real-time data fusion systems.
- To simulate an artificial AI-AI crisis scenario to discover the dynamics of escalation during compressed decision making.
- To determine the structural locations of inaccuracies in the AI-based preemptive reasoning, such as signal misinterpretation, automation bias, sensitivity to thresholds, and the eradication of diplomatic pause.
- To suggest strategic and policy mechanisms so as to maintain deterrence and reduce the risk of accidental escalation.

## 3: Methodology:

The proposed study uses a qualitative, analytical and scenario-based research approach to explore how why AI-enabled and high-speed decision-making can increase the risk of strategic errors in pre-emptive strike settings. Autonomous military systems can still only be experimented with directly by the government on operational secrecy and ethics; this limits the present state of this research to analysis based on historical background and, more structured, scenario modeling, but not quantitative experimentation. Principles of Cold War crises,

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

especially the false-alarm situation, near-miss crises provide empirical evidence of the role of timeline compression, ambiguity of intelligence and restraint of humans to influence crisis outcomes. These case studies show that even systems operated by humans were susceptible to misunderstanding through time pressure implying that even Artificial Intelligence-based acceleration can further escalate risks (Aritzis, F. 2025). Based on these lessons, the study creates a sequential AI-versus-AI crisis simulation that charts the course of detection, algorithmic threat measurement, intent definition, authorization levels, and response action. This formalized modeling methodology makes it easier to identify possible sources of automation bias, misclassification of data, and algorithmic overconfidence, and unintentional first-move escalation.

The analysis relies on pre-emptive strike theory, deterrence theory, and stability-instability paradox to provide the theoretical basis of the analysis, thus placing AI-mediated decision-making in the context of the prior strategic thought. The pre-emptive strike theory explains ways in which the perceived vulnerability may encourage swift AI-aided response; the deterrence theory explains how the credibility and signaling dynamics can be changed by the speed of algorithms; and the stability-instability paradox, the existence of stability at the strategic level that can be combined with instability at the tactical level, especially in circumstances where autonomous systems may act without much human intervention. Sources of data consist of declassified historical documents, official military doctrine related to AI-based command-and-control systems, and scholarly analyses addressing the topic of automation bias and the escalation procedures and new defense technologies (Kayode, B., Adebola, N. T., & Akerele, S. 2025).
The methodology is based on qualitative content analysis to identify recurrent patterns of decisions during the past crises, sequential scenario mapping to visualize escalation paths in AI-mediated interaction, and a framework of risk evaluation to estimate the technical failure, misinterpretation, cyber interference, and probability of overall escalation. The methodology suffers limitations in the form of relying on modeled assumptions and the fast developments in AI, however, triangulation based on historical evidence, the analysis of doctrines, and theoretical interpretation would improve the level of analytical rigor and applicability. The study provides a well-rounded and strategically sound evaluation of miscalculation risks that AI may introduce in the modern warfare through a mix of empirical precedent and future-oriented simulation.

## 4: Literature Review:

### 4.1: Preemptive Strike Theory

Preemptive strike theory is one of the foundations of the strategic research, foregrounding an offensive operation carried out to destroy the threat that is about to materialize before it can take place (Gray, C. S. 2007). Such action is rational in the eyes of decision-makers who anticipate disproportional losses or an existential threat when such delay is involved. Classic cases, such as the 1967 Six-Day war of Israel and cold-war nuclear one-upmanship, highlight the importance of perception of threat and fidelity of intelligence and timing in the application of preemptive reason. However, preemptive approaches are also fraught with the risk of wrong calculation, since any partial or distorted information will trigger unwarranted war.

This calculus is reinvented in AI integration. Autonomous decision-making systems are evaluating large amounts of data in real-time, measuring probabilities, and prescribing, or implementing, preemptive actions within minutes or even seconds. As much as these capabilities would speed up the reply, they would also limit the chances of a thoughtful response, people making decisions and diplomacy. As a result, the speed at which decisions

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

are made by AI can increase susceptibility to misunderstanding and unintentional heightening, especially when high-tension situations are prevailed by time-related indicators.

### 4.2: The Deterrence Theory and Crisis Stability

The deterrence theory offers a supplementary approach to the relationship between preemptive strike and the quick decision making. Deterrence works to make the opponents realize that the stakes of aggression are greater than the possible benefits, usually with the plausible worry of retaliation. Crisis stability consequently prevails when no one is encouraged to preempt on the other as the threat of escalation in a short period of time intimidates the advantages. The introduction of AI is a threat to the classical deterrence theories. Robotic systems that can react in almost real time squeeze decision timelines, which creates what researchers refer to as flash crisis situations.

This paradox of stability-instability has a special relevance to AI (Mirza, M. N., & Mujahid, H. A. 2025). AI-related technologies can increase the chances of strategic deterrence due to increased detection speed and precision targeting, but at the same time promote provocations of the lower level or miscalculations when the decision-making process becomes faster than the human reaction time. Researchers believe that AI creates new instabilities in the stability of crises, whereby opponents will interpret an algorithmic behavior as an act of aggression and induce unintended escalation (Boratov, S. 2025).

## 5: Automation and Decision-Making Models

The OODA (Observe-Orient-Decide-Act) is the conceptual framework that is often used in military decision making (Bryant, D. J. 2006). Conventional OODA loops that are human-centric are based on the iterative observation, contextual orientation, and deliberative decision-making. By automating data fusion, action selection and observation, AI integration will reduce the time these cycles take. As an example, autonomous targeting and early-warning systems are able to identify the launch of missiles, categorize the type of threat, and suggest counter-strike in several seconds.

Automation has its benefits as well as dangers. On the one hand, AI can lower the latency, alleviate human cognitive biases, and enable the fast coordination of assets that are distributed geographically. On the downside, algorithmic mistakes, false results, and misunderstanding of unclear indicators may cause improper reactions. Studies on automation bias show that human operators can over trust AI suggestions even in instances where outputs are inaccurate thus increasing chances of miscalculation (Hoff, K. A., & Bashir, M. 2015).

## 6: Stability-Instability Dilemma in AI-based Warfare

The stability-instability paradox argues that stability at the very top of the hierarchy may create conflict instability at the very bottom of the hierarchy. This is made difficult by AI, which hastens strategic and tactical decision-making. Though autonomous systems can prevent a significant attack using their precision and quick reaction, they can also misunderstand a minor provocation as something serious and can cause disproportionate reactions (Rastogi, A. 2020).

For instance, AI-enabled early-warning systems may be classified as threats from ambiguous radar signature as hostile launches and thereby precipitate preemptive alerts or automated counter-measures. With such automated reactions, the periods of time needed to evaluate humans are reduced, and this practically circumvents the customary diplomatic and deliberative protections. According to scholars, this undermines the so-called cushion of reflection that in the past enabled human actors to be restrained in case of a crisis.

## 7: Lessons of History in Thus Applied to AI

The importance of human judgment during crisis management is critical as witnessed in the past. The Cuban Missile Crisis showed that it is possible to avoid nuclear escalation using communication, interpretation of intentions and response that is measured.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

The threat of disastrous miscalculation in the speedy automated signaling told of catastrophe in case of a misinterpretation of the signals was further proven by the case of cold-war false alarm, e.g., the 1983 Soviet early-warning malfunction (Arbatov, A. 2017).

AI presents a similar challenge, namely, that speed advantage can be turned into a liability. Algorithms do not have the fine-tuning and ethical consideration of human actors, and therefore are likely to overestimate the threat or misunderstand unclear data. One of the strategies to mitigate AI integration is often suggested to be the integration with human oversight, so-called human-in-the-loop models, but the question of trade-offs between speed, accuracy, and human control continues to be discussed.

## 8: Past Precedence of Preemptive Logic

The consequences of preemptive strike logic with AI involvement require a historical crisis analysis based on which immediate decision-making and human acquaintance played a key role. The teachings based on such precedents provide a comparative structure with which to analyze the risk of magnifying the danger of miscalculation that may accompany accelerated AI decisions.

### 8.1: Cuban Missile Crisis (1962)

Cuban Missile Crisis is considered one of the most scrutinized examples of preemptive logic of high stakes in the modern history. In October 1962, the Soviet nuclear missiles were found in Cuba, and this led to a 13-day standoff between the United States and the Soviet Union (HARIHARAN, L. 2008). The leadership of the U.S. had the decision of either preemptive strikes that are instant and targeted or diplomacy that is delicate to make. Threats were imminent as indicated by intelligence reports but wrong interpretations would have turned the whole scenario into a nuclear war. Human forbearance was the determining factor. President John F. Kennedy and the people working with him used the deliberation, backchannel diplomacy, and strict evaluation of possible results. The decision to use a naval blockade as an alternative to a preemptive strike was an indication that the reckless move would lead to disastrous countermeasures. This crisis demonstrates that deliberation time, checking of intelligence and subtle interpretation of signals are important stabilizers-machineries that can be undermined in AI-faster settings.

### 8.2. False Alarm Incidents during the Cold War.

Automated early warning systems of the cold war brought new risks of preemptive miscalculation. Two striking examples reflect the possible outcomes of the fast, technological mediated decision-making: 1983 Soviet Early Warning System Incident: On September 26, 1983 Soviet satellite warning system falsely indicated that the U.S. rockets were heading the way. The lieutenant colonel who was in charge of alert reporting, Lieutenant Colonel Stanislav Petrov, treated it as a false alarm and he failed to escalate the alert to a retaliatory launch (Lewis, P., Williams, H., Pelopidas, B., & Aghlani, S. 2014). His move prevented a possible nuclear trade.

1979 NORAD Exercise False Alarm: An exercise at North American Aerospace Defense Command (NORAD) was mistaken with a Soviet attack. Quick response measures came close to alerting of a nuclear war before manual verification detected the mistake (Leuprecht, C., Sokolsky, J. J., & Hughes, T. 2018).

These cases show that the speed of technology is not the only goal that can ensure security. Humanity and judgment served as highly important stabilizers and did not allow automated warnings to develop into strategic disasters. On the contrary, AI systems might not have this judgment and contextual understanding thus increasing the chances of miscalculation in similar situations in the future.

### 8.3. There are some lessons that can be learned in Human Restraint in Crises.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

In all these historical examples, the same theme can be identified: human judgment and restraint can be considered as the key stabilizing elements in preemptive crises. Key lessons include:

- Checking of Threats: Preemptive action must have several levels of checkpoint to differentiate between real and false threats. Even with correct systems, the computerized systems can be misleading on the ambiguous signals.
- Deliberative Pause: There is time to think about other options, use consultants, and evaluate the possible outcomes; the less risky it is to the escalation. The decision-making process made by AI can circumvent these precautions.
- Contextual Awareness: Human beings are able to combine political, diplomatic, and moral factors that are currently unreachable when AI algorithms are involved. The knowledge plays a critical role in avoiding unintended growth.

These lessons imply that AI is capable of increasing the speed of detection and processing, but without the finer human control, these systems present some vulnerabilities. Unless necessary precautions are taken, preemptive reasoning on AI-based systems may turn small anomalies into an outright crisis.

## 9. Implications of AI Integration

The past experiences offer a background on the possible effect of AI on crisis dynamics. The experience of the Cuban Missile Crisis and the Cold War incidences has shown that speed as a benefit is not synonymous with strategic stability. The artificial intelligence systems used to shorten decision cycles should include the mechanisms of human intervention, multi-tiered verification, and context-driven analysis to reduce the risks of the unintentional escalation. Combining the lessons of history and AI-powered command and control systems can help expose the irony of high-speed decision-making: quicker response can not only discourage threatening actions but also encourage them. It is crucial to acknowledge this duality when policymakers and military planners have to work out the doctrine, protocols, and protection in the era of AI-based warfare

## 10. AI in Modern Command & Control

The current paper discusses the concept of implementing Artificial Intelligence (AI) into modern military command and control (C2) systems, where it is assumed that such implementation is a revolutionary change in the way of strategically positioning and the ability to conduct operations. The AI technologies such as autonomous targeting systems, early warning platforms, and data fusion networks can allow detecting, interpreting, and decision-making much faster than with the traditional human-centered processes. Although these improvements increase the efficiency of operations, they simultaneously provide increased risks of miscalculation and undesired increase especially in the case of preemptive strikes.

### 10.1 Multi-Core Autonomous Targeting Systems

The autonomous targeting is among the most impactful AI applications in military activities. These systems use machine-learning algorithms to detect, classify, and rank possible targets in real time using sensor data (Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. 2018). By automating, the decision time can be significantly shortened and preemptive attacks can be carried out in a matter of seconds. As an example, the current unmanned aerial vehicles (UAVs) and missile defense systems also incorporate AI to automatically track the trajectories of adversaries, determine the likelihood of threat, and deploy the countermeasures. As much as automation enhances response time and minimizes the risk of human error in high-tempo engagements, the issue is that algorithms have an over-exerting nature, and therefore, unclear or otherwise incomplete data can be interpreted as a threat to be addressed. The result of such misinterpretation may be the unwanted involvement, particularly during the crisis situation when the process of real-time verification is limited.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

### 10.2. Early Warning Systems

Early warning systems having artificial intelligence combine satellite constellation, radar networks and electronic intelligence in order to provide near real-time situational awareness (Solovyeva, A., & Hynek, N. 2023). These systems, which are designed to identify the launch of missiles, the movements of aircrafts or naval actions, can offer actionable intelligence to the decision-makers faster than traditional monitoring. Early warning is enhanced by AI which is fast and precise thereby increasing deterrence by giving credible and timely warnings. However, the historical events including the 1983 Soviet false alarm help to realize that even slight mistakes in recognizing a threat may lead to disastrous growth. Artificial intelligence software is vulnerable to false positives, sensor failures or adversarial attacks, which could lead to premature strikes without human intervention.

### 10.3. Data Fusion Systems

One of the fundamental AI functions of the current C2 networks is data fusion. The AI algorithms combine a unified operational picture by taking into consideration a wide array of heterogeneous data such as signals intelligence, satellite images, and sensors in the battlefield. This combination lowers the uncertainty, increases predictability, and allows assessing a number of simultaneous threats quickly. Nevertheless, the sophistication of AI-data fusion also creates such vulnerabilities as the misalignment of datasets, the over fitting of predictive models, and bias in training data, which all can generate false threat prioritization. One misclassification in high stakes situations can increase tension throughout the operations spectrum, and this is the reason why layered verification and human intervention are required.

### 10.4 Speed of OODA Loop Compression and Decision

The AI essentially shrunk the Observe-Orient-Decide-Act (OODA) loop, a pillar of the military decision making theory. Conventional human processes comprise repetitiveness of observation, contextual orientation, deliberation and action. This Vicious Cycle is expedited by AI, which tracks several data streams in parallel, orienting on pre-designed threat-related characteristics, and prescribing (or even acting) on responses independently. Even though expedited OODA cycles improve reaction time and operational effectiveness, it can limit chances to have diplomatic pause, moral scrutiny and intelligence cross-checking. As a result, the speed increases the likelihood of error, especially when opponents will interpret automated behavior to be hostile, thus triggering spirals of escalation (Collins, R. 2012).

## 11. Threats AI errors, False Positives and Automation Bias

With the introduction of AI into C2 systems, there are new types of risk:

- Technical Errors: Sensors and malfunctioning software or erroneous calibration can result in faulty threat determination.
- False Positives: An over-sensitive detector can fail to recognize harmless processes as hostile and, therefore, respond with adverse actions (Holden, S. 2025).
- Automation Bias: With AI-generated information, human operators can be overly confident in the accuracy of the results and fail to critically evaluate them, thus, reducing human decision-making in times of crisis (Kovari, A. 2024).

Altogether, these threats prove that speedy decision-making is not necessarily connected with increased security. AI systems increase speed and situational awareness but have the potential to increase miscalculation unless it is reduced by effective human controls.

### 10.5 Strategic Stability implication

Artificial intelligence in C2 systems transforms the conventional deterrence and escalation relationships. On the one hand, swift identification and independent reaction abilities enhance deterrence by making it apparent that there are valid defense or retaliatory capabilities. Conversely, AI can hasten the process of crisis response and increase the possibility of unintentional escalation and shortening the time available to implement diplomatic actions

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

(Johnson, J. 2023). Such a duality contributes to the dire need to structure human-in-the-loop processes, strict verification rules, and ethical limitations that must be introduced to AI-enabled C2. Stabilizing AI presence strategically in an AI-armed world thus requires a trade-off between the operation advantages of speed and countermeasures which ensure unintended escalation is avoided.

## 11. Hypothetical Crisis Scenario (AI vs. AI)

To analyze the dangers of AI-based preemptive reasoning, this paper will establish a hypothetical scenario that involves two autonomous systems in a crisis. The situation shows the way in which autonomous systems, quick interpretation of threats, and algorithmic retaliation may interact in a fast military setting, and, therefore, may potentially raise the risk of a miscalculation.

### 11.1. Scenario Overview

As an example, take two warring states, State A and State B, with fully developed AI systems of command and control (C2), autonomous targeting platforms, and built in early warning systems. There is a local conflict on disputed airspace and maritime traffic. Both artificial intelligence systems constantly track the satellite, radar, and electronic intelligence, analyzing signals to determine threats. When the State A identifies an unknown launch of missiles in the vicinity of its borders, the crisis has begun. Its artificial intelligence system determines the launch as a likely hostile move by State B with an estimate of a high probability of immediate attack. At the same time, the AI of State B interprets the defensive actions of State A as an act of offensive escalation. It is this situation, therefore, that provokes an algorithmic feedback loop, in which quick detection and automated threat analysis reduce decision-making times to seconds.

### 11.2. Stepwise Escalation Pathway

- Detection of Threat: To identify what may be considered Cape Gray activities, AI systems analyze multiple data streams, including satellite photographs, radar scan returns, and intercepted communications, to do so.
- Rapid Threat Interpretation: Both AI assess intent based on the parameters that are pre-programmed, past information, and probabilistic models. The system can treat ambiguous signals as hostile since the system does not tolerate uncertainty thus leading to pre-emptive alert programs (Horowitz, 2016; Lin, 2016).
- Algorithmic Retaliation: The AI proposes or automatically retaliates, according to perceived threats, with countermeasures (e.g. missile interception or electronic jamming). The AI of the system perceives the rapid response of each system as an escalation, and responds further in defense or retaliation.
- Escalation Spiral: When OODA loop is compressed, the threat detection, interpretation, and action will result in a feedback loop without human intervention. Small…

### 11.3. Key Dynamics

- Ambiguity and Misinterpretation: The AI systems do not provide a finer perspective as to intent and context. An action that is meant to be defensive can be construed as offensive and creates a vicious cycle of self-reinforcement.
- Lack of Diplomatic Pause: The classical crises usually have room of negotiation and confirmation. Algorithms in AI-driven interactions can remove this window and, consequently, the likelihood of engaging in a pre-emptive interaction increases.
- Algorithmic Reciprocity: The reactions of each AI system to the other system in an automated fashion create a cycle of algorithmic retaliation that can be faster than the human decision makers.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

- Threshold Sensitivity: AI decision threshold, which works on fast-response rates can cause pre-emptive action in case of minor anomaly or sensor failures, which introduces the risk of false positives that can lead to real-life escalation.

### 11.4. What it means on Strategic Stability.

This hypothetical situation proves that the acceleration of AI may lead to the destabilization of the crisis unexpectedly. Although AI increases the speed of detection and decisions, it increases the chances of miscalculation:

- Even minor mistakes are exacerbated with a quick reaction of algorithms.
- Lack of human judgment will eliminate chances of contextualized restraint.
- A wrong understanding of the automated activities can turn ordinary defense to a perceived assault.

The scenario demonstrates the artificiality of AI-enabled speed by modeling such interactions: the systems that are intended to enhance security and responsiveness can at the same time pose an additional threat of unintended escalation and strategic miscalculation

### 11.5. Policy and Doctrine Relevancy.

Analysis through scenarios suggests that safeguards should be used, such as:

- Human-in-the-loop or human-on-the-loop surveillance of pre-emptive choices.
- Multi-layer information validation of sensor information and classes of threats.
- Automatism constraints, which put confirmation first and then retaliate automatically.
- Procedures involving exchanges between the enemy to ensure that they do not wrongly interpret automated actions.

The measures will strengthen the strategic stability and take advantage of operational benefits of AI-enhanced decision-making.

## 12. Points of Miscalculation

The introduction of AI in military decision-making process is both accelerating preemptive logic and also creates unique vulnerabilities at the same time. Any mistake can appear at various points of a crisis, during the detection of threats, up to autonomous response, but this is not always perceived by a person immediately. The understanding of these arguments is critical to the reduction of the risks of inadvertent escalation.

### 12.1. Signal Misinterpretation

The AI systems are highly dependent on sensor data, satellite images, radar information, and electronic data to evaluate threats (Horowitz, M. C., Allen, G. C., Saravalle, E., Cho, A., Frederick, K., & Scharre, P. 2022). Nonetheless, the unclear or partial signals are likely to be misunderstood as hostilities. As an example, a regular military drill, missile test or electronic anomaly can cause an automatic threat notification. The absence of a refined contextual understanding that allows interpretive restraint is something AI does not have. An example of the 1983 Soviet early-warning false alarm proves that any slight mistake in classification of the threat can nearly trigger catastrophic escalation. In an AI-driven system, these misunderstandings may spread now very quickly, and the algorithmic countermeasures may be activated even before humans can take action.

### 12.2. Escalation Without Intent

Unintended escalation is another very important area of miscalculation. In situations where an auto-interpretation of defensive maneuvers or reconnaissance activity is perceived as an aggressive move by the AI systems, it might begin to preempt or retaliate. This effect can be increased by the reciprocal character of AI-enabled systems, whereby each party responds to automated activity in near real-time, which forms a feedback loop of escalation. Even smaller actions in the operation process can escalate to bigger conflicts which highlights the irony that the quicker the response, the more the instability.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

### 12.3. Lack of Diplomatic Pause

Diplomatic pause is a hallmark characteristic of the human-led crisis management, meaning the period of time in which the verification, consultation, and negotiation can occur. This delay is dramatically decreased or removed by AI acceleration, squeezing down the decision-making time to seconds (Pokhriyal, N., & Koebe, T. 2023). The lack of deliberative time will inhibit the possibilities of clarifying intentions, indicating that things will be held back, or that ambiguity can be resolved by non-military means. Any little misunderstanding or false alarm can therefore shoot into irreparable preemptive steps before diplomatic or human corrective action can come in.

### 12.4. Automation Bias and Over-reliance on Humans

Automation bias, the desire to rely on AI results more than they would their intuition, can also be a contributor to errors in judgment when humans still participate in the decision-making process (Pokorny, L. 2025). The AI may be used to make judgments that are incorrect by the operator without analyzing conflicting intelligence or other interpretations. This over-reliance reduces critical assessment, which is more or less a form of procedural obligation, instead of a substantive protection.

## 13. Policy Directions

The inclusion of artificial intelligence in the military command and control requires serious protection measures to prevent sudden errors and unintentional progressions.

- Mandatory Human Control: Human approval of strategic and high-risk decisions, especially decisions dealing with nuclear forces, should be explicitly mandated; Artificial Intelligence should be subordinate to human judgment and experience in politics and military affairs and not an alternative.
- Multi-Source Verification: The possibility of false alarms and misclassification should be reduced by assuming that automated response mechanisms are based on the cross-checked intelligence input.
- Well-defined Operational Bounds: Autonomous systems should be able to work within a set engagement limits and also should not be able to change autonomously between detection and strike.
- Transparency Mechanisms: Crisis communication mechanisms and signaling mechanisms must be structured in order to reduce misinterpretation between adversaries.
- Accountability and Auditability: Artificial intelligence systems should maintain traceable decision histories so that they can be accountable and it will be easy to review them after crisis.
- International Norms: The states ought to create collaborative systems that limit completely autonomous strategic retaliation and promote responsible use of AI.

## Conclusion

The introduction of Artificial Intelligence (AI) into the military command and control architectures is a radical change in the dynamics of operations that underlie the decision-making process of preemptive strikes. The current investigation tested the hypothesis that AI-driven acceleration increases the risk of high-level strategy errors during crises in the military. The analysis supports the argument that velocity provides operational competitiveness, but at the same time, it creates increased
instability in an unrestrained state by integrating past case studies, strategic doctrine, and simulations based on specific scenarios.

The history of agitation shows, that man has time and again avoided final disaster by the judgment of man, by a verification which supports, by self-control. On the other

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

hand, the fact that AI compresses deliberative timelines cuts off the space of contextual inference and political deliberation. Even though augmented detection and fast response can strengthen deterrence under a few circumstances, it also increases the consequences of false positives, inaccuracies in data, and misclassification of algorithms.

Scenario simulations also indicate that fast threat evaluations and retaliation automated in the context of AI may breed escalation spiral out of planned political control. Where machines operate at their speed, equivocal signals have the potential of triggering involuntary reflexes that are against the will of human decision-makers. Therefore, acceleration is not an inherently safe aspect because in the event of a lack of sufficient protective measures; it increases the probability of accidental escalation.

The paper also finds that the strategic implications of AI depend not only on the technological capability alone but also on the governance structures. The human control, multistage inspection, high-terms of engagement, and moral limitations are still invaluable tools in the risk reduction. The crisis stability is further supported by the international norms and the doctrinal recalibration.

Finally, AI does not stabilize or destabilize in itself. The influence on strategic equilibrium depends on the manner in which states incorporate it in controlled, accountable, and restraint-based systems. The technological speed should be brought up to pace with the institutional caution to prevent wrong judgments in the changing environment of AI-assisted warfare.

**References:**

Mueller, K. P., Castillo, J. J., Morgan, F. E., Pegahi, N., & Rosen, B. (2006). *Striking first: preemptive and preventive attack in US national security policy* (Vol. 375). Rand Corporation.

Chang, B. A. (2021). *Artificial Intelligence and the US-China Balance of Power* (Doctoral dissertation, Massachusetts Institute of Technology).

Kokoshin, A. A. (2007). *Nuclear Conflict in the Twenty-first Century*. Belfer Center for Science and International Affairs, John F. Kennedy School of Government, Harvard University.

Panda, A. (2025). *The New Nuclear Age: At the Precipice of Armageddon*. John Wiley & Sons.

Johnson, J. (2023). *AI and the bomb: Nuclear strategy and risk in the digital age*. Oxford University Press.

De Visser, E. J., Pak, R., & Shaw, T. H. (2018). From 'automation'to 'autonomy': the importance of trust repair in human–machine interaction. *Ergonomics*, *61*(10), 1409-1427.

Aritzis, F. (2025). The Impact of Artificial Intelligence on the Consumer Behavior of Greeks: Understanding the Awareness of AI Benefits in Their Consumer Journey and How Lack of Knowledge of the Concept of Artificial Intelligence Can Lead to Ignorance, Anxiety, and Fear. *The Impact of Artificial Intelligence on the Consumer Behavior of Greeks: Understanding the Awareness of AI Benefits in Their Consumer Journey and How Lack of Knowledge of the Concept of Artificial Intelligence Can Lead to Ignorance, Anxiety, and Fear*.

Mircheska, S. (2025). The Future of Military Leadership: How Artificial Intelligence is Changing Military Conflict. *Сигурност и отбрана*, (1), 104-113.

Perla Jr, H. (2005). *Revolutionary deterrence: the Sandinista response to Reagan's coercive policy against Nicaragua, lessons toward a theory of asymmetric conflict*. University of California, Los Angeles.

Bolan, C. J. (2009). *Risk in American foreign military interventions*. Georgetown University.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Danilovic, V. (2002). *When the stakes are high: Deterrence and conflict among major powers*. University of Michigan Press.

Mirza, M. N., & Mujahid, H. A. (2025). AI Arms Race And Strategic Stability Between India And Pakistan. *IPRI Journal*, *25*(01), 10-31945.

Johnson, J. (2023). *AI and the bomb: Nuclear strategy and risk in the digital age*. Oxford University Press.

De Spiegeleire, S., Maas, M., & Sweijs, T. (2017). *Artificial intelligence and the future of defense: strategic implications for small-and medium-sized force providers*. The Hague Centre for Strategic Studies.

Bryant, D. J. (2006). Rethinking OODA: Toward a modern cognitive framework of command decision making. *Military Psychology*, *18*(3), 183-206.

Johnson, J. (2023). *AI and the bomb: Nuclear strategy and risk in the digital age*. Oxford University Press.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors*, *57*(3), 407-434.

Rastogi, A. (2020). *Trust and Anti-Autonomy Modelling of Autonomous Systems* (Doctoral dissertation, North Dakota State University).

Writer, N. D., Ahmed, S., Bajema, N. E., Bendett, S., Chang, B. A., Creemers, R., ... & Weber, V. (2019). Artificial Intelligence, China, Russia, and the Global Order Technological, Political, Global, and Creative Perspectives.

Arbatov, A. (2017). Understanding the US–Russia Nuclear Schism. *Survival*, *59*(2), 33-66.

HARIHARAN, L. (2008). Cold War Negotiations: Berlin Crisis & Cuban Missile Crisis.

Lewis, P., Williams, H., Pelopidas, B., & Aghlani, S. (2014). *Too close for comfort* (Doctoral dissertation, Chatham House-Royal Institute of International Affairs).

Wilson, G. A. (2012). *NORAD and the Soviet Nuclear Threat: Canada's Secret Electronic Air War*. Dundurn.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *arXiv preprint arXiv:1802.07228*.

Solovyeva, A., & Hynek, N. (2023). When stigmatization does not work: over-securitization in efforts of the Campaign to Stop Killer Robots. *AI & SOCIETY*, *38*(6), 2547-2569.

Collins, R. (2012). C-escalation and D-escalation: A Theory of the Time-dynamics of Conflict. *American Sociological Review*, *77*(1), 1-20.

Maleki Varnosfaderani, S., & Forouzanfar, M. (2024). The role of AI in hospitals and clinics: transforming healthcare in the 21st century. *Bioengineering*, *11*(4), 337.

Holden, S. (2025). *Safeguarding Accountability in the UK: Whistleblowing as a Response to State Secrecy and Abuse of Power* (Doctoral dissertation, Manchester Metropolitan University).

Johnson, J. (2023). *AI and the bomb: Nuclear strategy and risk in the digital age*. Oxford University Press.

Horowitz, M. C., Allen, G. C., Saravalle, E., Cho, A., Frederick, K., & Scharre, P. (2022). *Artificial intelligence and international security*. Center for a New American Security..

Johnson, J. (2022). Delegating strategic decision-making to machines: Dr. Strangelove Redux?. *Journal of Strategic Studies*, *45*(3), 439-477.

Handfield, R., & Linton, T. (2017). *The LIVING supply chain: The evolving imperative of operating in real time*. John Wiley & Sons.

Pokorny, L. (2025). Human-AI Collaboration in High-Stakes Decision-Making Environments.