AI in Conflict Prevention

28 JUNE 2025

Muejgan Ekberli



In recent years, the development of artificial intelligence (AI) has led to a significant transformation across many fields, including governance, security, and humanitarian intervention (Pauwels, 2020; Mandarava, 2024). Once limited to the sphere of computer science and robotics, AI is now integrated into institutional frameworks and normative discourses surrounding international peace and conflict prevention (Pasupuleti, 2025). A crucial question emerges as AI progressively enters early warning systems, predictive analytics, and peacekeeping infrastructure: To what extent can machine learning and data-driven algorithms contribute to anticipating and preventing conflict, and to what extent might they instead introduce new uncertainties, biases, and surveillance practices?

Conflict prevention has traditionally relied on human-led diplomacy, social intelligence, and complex institutional coordination. These methods face systemic challenges in fragile or "rapidly changing environments" (Desai, 2020, p. 23). AI technologies, by contrast, promise real-time risk identification, enhanced pattern recognition, "anomaly detection and behavioral analysis" (Pauwels, 2020, p. 2). Such capabilities offer the potential to anticipate conflict dynamics before they escalate into violence, thereby enabling earlier and potentially more effective interventions (Mandarava, 2024; Kimaita & Irungu, 2024). However, these promises are embedded in critical uncertainties concerning data quality, algorithmic transparency, ethical accountability, and policy frameworks (Amen, 2024; Cheong, 2024).

Many initiatives underscore both the promise and complexity of using AI in fragile political contexts. For instance, the Global Conflict Risk Index (GCRI), developed by the Joint Research Center of the European Commission, applies random forest algorithms that support early warning and conflict prevention efforts (Ferri et al., 2019). The AI-enhanced GCRI uses structural risk indicators, e.g. "regime type and [...] GDP per capita" (Ferri et al., 2019, p. 4), to predict the probability and intensity of conflicts at national and subnational levels. Similarly, pilot projects in Kenya illustrate the feasibility of integrating AI into early warning and response systems (EWRS), where machine learning algorithms are trained to detect rising tensions "based on historical, current and emerging data" (Kimaita & Irungu, 2024, p. 2332).

The following chapters explore the dual character of AI in conflict prevention. The next section outlines the main potentials of AI while the final section critically discusses the challenges and ethical considerations.

2. Potentials of AI in Conflict Prevention

The potential of AI in conflict prevention lies in its capacity to analyze complex and large-scale datasets and generate and offer insights faster and more accurately than traditional methods (Pauwels, 2020; Mandarava, 2024). Early warning systems represent one of the most advanced applications, where AI supports the identification of conflict drivers through data analysis of trends, geographic distributions and linguistic dimensions (Mandarava, 2024). Once sufficiently trained, machine learning models can identify irregularities in communication, shifts in economic indicator, or signs of group mobilization, enabling the early detection of emerging risks before they become apparent to human analysts (Ferri et al., 2019; Kimaita & Irungu, 2024).

In Kenya, for instance, AI-driven early warning and response systems are able to improve situational awareness by combining social media analysis with historical conflict records and live geospatial data (Kimaita & Irungu, 2024). The system enables more responsive and localized conflict mediation measures. According to Mandarava (2024), social media platforms play

a critical role in such processes, allowing analysts to trace sentiment shifts, rising dissatisfaction, or coordinated disinformation campaigns. Consequently, real-time monitoring of social media keywords and emotional dynamics can serve as early indicators of potential instability (Mandarava, 2024).

Furthermore, AI improves conflict mapping and key actor monitoring by using natural language processing and computer vision to analyze news media, social media platforms, and satellite imagery for indicators such as military mobilization, hate speech spread, or critical infrastructure damage (Pauwels, 2020; Abedin et al., 2025).

3. Case Studies: Real-World Applications of AI in Conflict Prevention

The application of AI in conflict prevention is no longer a theoretical concept but a growing reality across different regions and conflict settings. Several real-world case studies highlight how AI-driven systems have contributed to early warning and civilian protection.

One of the advanced projects is the Violence Early-Warning System (ViEWS) (Hegre et al., 2019). This system "produces monthly forecasts at the country and subnational level for 36 months into the future and all three UCDP types of organized violence: state-based conflict, non-state conflict, and one-sided violence in Africa" (Hegre et al., 2019, p. 155). The ViEWS framework is composed of multiple constituent models that draw upon "decades of quantitative peace and conflict research" (Hegre et al., 2019, p. 156). Its predictive power is valuable for international organizations and NGOs seeking to allocate resources and design preventive interventions.

In the context of active conflict, the Sentry system, developed by Hala Systems in Syria, stands out as an example of a tactical-level AI application (Hale et al., 2023). This system uses acoustic sensors, human input, and machine learning algorithms to detect and predict airstrikes in real-time. Once an imminent attack is identified, warnings are disseminated via mobile apps and social media to alert civilians and humanitarian responders. According to data published, "when the EWS consisted of sirens and social media messages, indicated that overall mortality was reduced by 20% – 30% when timely warnings are available" (Hale et al., 2023, p. 22). The Sentry system underscores the potential of AI to enhance civilian protection by providing timely, localized alerts in high-risk environments.

In the realm of conflict simulation and decision-support tools, the CulturePulse platform developed in partnership with the United Nations represents a novel use of AI for peacebuilding (Gilbert, 2023). This initiative applies a digital twin approach to simulate the sociopolitical conflicts, such as the Israeli-Palestinian conflict. By modeling over "80 categories to each "agent," including traits like anger, anxiety, personality, morality, family, friends, finances, inclusivity, racism, and hate speech" (Gilbert, 2023), Cul-

turePulse enables policymakers to test the potential outcomes of different interventions in a virtual environment.

Another example is the peace process in Yemen, where researchers use machine learning tools to analyze transcripts of mediation dialogues (Arana-Catania et al., 2022). By identifying changes in language, the system provides insights into the evolving positions of the conflicting parties (Arana-Catania et al., 2022). This enables mediators to adjust their strategies and anticipate potential negotiation deadlocks more effectively. The project illustrates how AI can support the intangible aspects of peace negotiations by uncovering communication patterns and facilitating more informed mediation strategies.

4. Challenges and Ethical Considerations in the Use of AI for Conflict Prevention

The increasing role of AI in conflict prevention offers great promise, but it also brings significant challenges and ethical considerations. One key challenge is bias in AI systems (Min, 2023). AI algorithms are trained on data, and if this data reflects existing societal biases, the AI system will likely reproduce and even amplify these biases (Min, 2023; Ntoutsi et al., 2020). For instance, if a risk assessment tool used in criminal justice is trained on historical criminal records that contain racial disparities, it may unfairly over-predict the risk of recidivism for certain racial groups (Min, 2023). This can result in discriminatory outcomes and further reinforce existing inequalities within the justice system (Min, 2023; Ntoutsi et al., 2020). Therefore, according to Cheong (2024), "meaningful AI accountability requires grappling with power imbalances between AI developers and those affected by their systems" (p. 8). Thus, it is crucial to include marginalized groups "who are most at risk of AI harms" (Cheong, 2024, p. 8) in policymaking.

Another challenge is ensuring the responsible and ethical collection and use of data (Min, 2023). AI systems often rely on vast amounts of data, including personal information, to identify patterns and make predictions. However, the collection and use of this data can raise privacy concerns, especially for marginalized communities who may be disproportionately affected by privacy violations (Min, 2023). Min (2023) also states that it is important to implement privacy-preserving techniques and ethical data labeling practices to protect sensitive attributes and avoid bias and stereotypes. Building on this, according to Pauwels (2020), "[c]onflict prevention actors must understand the computational techniques on which they rely and the data sets in use, particularly how data is collected and the biases those data sets may represent" (p. 16).

Moreover, algorithmic fairness is also a critical ethical consideration (Min, 2023). It refers to designing AI algorithms that produce equitable outcomes across different demographic groups. However, achieving both

fairness and accuracy can be challenging, as fairness constraints may restrict the ability of the model to make accurate predictions (Min, 2023) which makes it more difficult to accurately predict conflicts. Thus, according to Min (2023), it is crucial to find a balance between fairness and accuracy to ensure that AI systems deliver equitable outcomes without undermining their overall performance. This can be achieved by integrating fairness constraints into the objective function of the model during training and by applying regularization techniques that penalize biased predictions, thereby promoting algorithmic fairness.

Furthermore, structural imbalances in the development and accessibility of AI technologies contribute to deepening existing geopolitical inequalities. While AI is sometimes portrayed as a democratizing force given the global accessibility of tools like ChatGPT, "AI development is still concentrated in a handful of states and companies" (Srivastava et al., 2024, p. 14). These developments raise concerns about technological dependency, data sovereignty, and epistemic inequality (Srivastava et al., 2024). Consequently, disparities in access to AI technologies may lead to conflict prevention strategies that disproportionately reflect the interests of powerful nations and corporate actors.

Finally, the integration of AI into surveillance and behavioral prediction systems may infringe on privacy rights and civil liberties (International Committee of the Red Cross, 2019), particularly when used by authoritarian regimes or unregulated private actors (Pauwels, 2020). Therefore, in the International Review of the Red Cross (2020), the International Committee of the Red Cross (ICRC) emphasizes the importance of maintaining human control and ethical oversight in AI applications during armed conflict and humanitarian operations. Mandarava (2024) further notes that the growing role of cyberspace in conflict settings requires robust international legal standards for digital governance.

Furthermore, the potential for AI to displace humans from decision-making processes raises further ethical questions (Reder & Koska, 2024). While AI can enhance decision-making by providing data-driven insights, it is important to maintain human oversight and judgment. Thus, determining right from wrong is a fundamentally human responsibility and should remain so in the future.

In conclusion, while AI holds transformative potential for conflict prevention, its benefits must be weighed against significant ethical, political, and technical constraints. A human-centered, participatory, and transparent governance framework is essential to harness the strengths of AI without compromising fundamental rights or undermining local agencies. In addition, bias in AI development also must be addressed to ensure fairness and prevent unjust outcomes.

5. Conclusion

From early warning systems to peace negotiation support, AI is already shaping how we identify risks, protect civilians, and design interventions. These new tools allow for faster analysis, broader data integration, and, in some cases, more timely responses than traditional, human-led approaches.

However, with these advances come important questions: How do we ensure that AI supports rather than replaces human judgment? Can we trust these systems to be fair, especially in fragile political contexts where data is often incomplete or biased? And who ultimately controls the infrastructures behind these technologies? What emerges is a dual responsibility which is to embrace the opportunities AI provides while being critically aware of its limitations and risks.

In summary, the integration of AI into conflict prevention requires more than technological innovation. It demands inclusive governance frameworks, ethical reflection, and sustained interdisciplinary collaboration. Ensuring that AI serves as a tool for peace, rather than a new source of inequality or harm, depends on how carefully and deliberately these systems are designed, implemented, and regulated. To ensure that AI serves as a force for peace rather than an instrument of surveillance or control, future efforts must center on transparent, participatory, and context-sensitive implementation.

References

Abedin, A., Bais, A., Buntain, C., Courchesne, L., McQuinn, B., Taylor, M. E., & Ullah, M. (2025). A call to arms: Automated methods for identifying weapons in social media analysis of conflict zones. Computer Vision and Pattern Recognition (cs.CV); Human-Computer Interaction (cs.HC), 306–314. https://doi.org/10.1145/3717867.3717886

Amen, M. A. (2024). AI-Driven Sustainable Habitat Design: key policy frameworks and ethical safeguards. Smart Design Policies, 1(1), 23–32. https://doi.org/10.38027/smart-v1n1-4

Arana-Catania, M., Van Lier, F., & Procter, R. (2022). Supporting peace negotiations in the Yemen war through machine learning. Data & Policy, 4. https://doi.org/10.1017/dap.2022.19

Cheong, B. C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. Frontiers in Human Dynamics, 6. https://doi.org/10.3389/fhumd.2024.1421273

Desai, H. (2020). Conflict prevention in fragile contexts. OECD Development Co-operation Working Papers. https://doi.org/10.1787/041cbaf0-en

Ferri, F., Yannick, D., Michail, P., Marie-Sophie, V. D., Kathrin, B., & Stamatia, H. (2019). The Global Conflict Risk Index: Artificial Intelligence for Conflict Prevention. JRC Technical Reports. https://doi.org/10.2760/004232

Gilbert, D. (2023). The UN hired an AI company to untangle the Israeli-Palestinian crisis. WIRED. https://www.wired.com/story/culturepulse-ai-israeli-palestinian-crisis/?

Hale, S., Poulson, C., Winslow, E., Triple Line, Mahrousa, M. B., Alhayes, M., & Jouri. (2023). Creating Hope in Conflict: A Humanitarian Grand Challenge contracted Triple Line Consulting to deliver a series of outcome and value for money case studies for a sample of funded innovations.

Hegre, H., Allansson, M., Basedau, M., Colaresi, M., Croicu, M., Fjelde, H., Hoyles, F., Hultman, L., Högbladh, S., Jansen, R., Mouhleb, N., Muhammad, S. A., Nilsson, D., Nygård, H. M., Olafsdottir, G., Petrova, K., Randahl, D., Rød, E. G., Schneider, G., Vestby, J. (2019). ViEWS: A political violence early-warning system. Journal of Peace Research, 56(2), 155–174. https://doi.org/10.1177/0022343319823860

International Committee of the Red Cross. (2019). Artificial intelligence and machine learning in armed conflict: A human-centred approach.

International Review of the Red Cross. (2020). Artificial intelligence and machine learning in armed Conflict: a human-centred approach. International Review of the Red Cross, 102(913), 463–479. https://doi.org/10.1017/s1816383120000454

Kimaita, S., & Irungu, E. J. (2024). New frontiers in conflict Prevention: Integrating artificial intelligence in early warning and response systems in Kenya. International Journal of Research and Innovation in Social Science, VIII(XI), 2331–2339. https://doi.org/10.47772/ijriss.2024.8110185

Mandarava, T. L. (2024). International Security Governance in the digital Era: Navigating challenges in conflict prevention, resolution, and disaster response. In Advances in Economics and Management Research FITED 2024 (Vols. 12–2024).

Min, A. (2023). Artificial Intelligence and Bias: Challenges, Implications, and Remedies. International Journal of Scientific Research, 2(11), 3808–3810. http://ijsr.internationaljournallabs.com/index.php/ijsr

Moshtagi, R., Cortez, J., & Sohn, A. (2025, May 27). AI and the Future of Conflict Resolution: How Can Artificial Intelligence Improve Peace Negotiations? The Belfer Center for Science and International Affairs. https://www.belfercenter.org/research-analysis/ai-and-future-conflict-resolution-how-can-artificial-intelligence-improve-peace?

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder‐Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., .Staab, S. (2020). Bias in data‐driven artificial intelligence systems—An introductory survey. Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, 10(3). https://doi.org/10.1002/widm.1356

Pasupuleti, M. K. (2025). AI's role in global stability, diplomacy, and peacebuilding. In International Journal of Academic and Industrial Research Innovations (pp. 1–11). https://doi.org/10.62311/nesx/32268

Pauwels, E. (2020). Artificial intelligence and data capture technologies in violence and conflict prevention. https://www.jstor.org/stable/resrep27551

Reder, M., & Koska, C. (2024). Künstliche Intelligenz und ethische Verantwortung. In Michael Reder, Christopher Koska (Ed.), Edition Moderne Postmoderne. transcript Verlag.

Srivastava, S., Bullock, J., Political Science, Purdue University, Bush School of Government & Public Service, Texas A&M University, Convergence Analysis, & Global Governance Institute. (2024). AI, global governance, and digital sovereignty.

DOI 10.69790/epis-rep-pcp-1/39133