Re-Programming Peace: AI and the Future of Conflict Resolution

Ruth Mathen

Manish Sharma

OCTOBER 2025

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs

About the Emerging Technology, Scientific Advancement, and Global Policy Program

Science and technology have long played a role in international affairs, but today they are more central to geopolitics than ever. The Program on Emerging Technology, Scientific Advancement, and Global Policy responds to an urgent need to advance scholarship and build a new generation of leaders fluent in both public policy and technology.

The Program conducts research, convenings, training, and public engagement. Our work includes:

- Technology and Geopolitics Fellowships for outstanding pre- and post-doctoral scholars

- Senior Fellowships to bring experienced scholars and practitioners into our scholarship, events, mentorship, and training

- Reports and articles

- Seminars, study groups, and other convenings

This White Paper is a product of the AI+X+G initiative, a new study group under the Emerging Technology, Scientific Advancement, and Global Policy Program. Through a series featuring prominent speakers and student-led discussions, the study group explores the intersection of AI and policy challenges and identifies potential policy and action plans.

About the Authors

Lead Authors

Ruth Mathen

Ruth Mathen is committed to strengthening international governance on transnational challenges such as migration, security, and peace, with a focus on building inclusive and just multilateral systems. She is a Master in Public Policy candidate at the Harvard Kennedy School, specializing in international and global affairs, and a Belfer Young Leader Fellow. Previously, Ruth led regional programs on migration, peacebuilding, and human rights at the Christian Conference of Asia, an ECOSOC-accredited international ecumenical organization representing over 65 million people across seventeen countries. In this role, she advanced cross-country initiatives, managed the organization's communications, and served as its liaison to the United Nations. She also serves with the World Council of Churches, where she is a youth advisor to the Central Committee and a core member of the Gender Justice Reference Group. Ruth holds a BA in economics and sociology from St. Xavier's College, Mumbai, and a Professional Diploma in sustainable development and human rights from the University for Peace.

Manish Sharma

Manish Sharma recently completed the Mid-Career Master in Public Administration at the Harvard Kennedy School, where he focused on leveraging AI for business and social impact. Prior to Harvard, he spent over twelve years in India's education sector, working to improve learning outcomes in public schools and expand higher education access for marginalized communities. He co-founded and scaled a nonprofit that enabled thousands of underprivileged students to secure scholarships and mentorship. He also co-led a multistate consortium that transformed public scholarship delivery systems through the use of data and technology. Previously, Manish served as Strategic Initiatives Lead at one of India's largest education nonprofits, where he developed long-term strategies for major educational reform programs and improved learning outcomes. He also holds a Bachelor of Science in electronics and communication engineering.

Contributors

Josh Cortez

Josh Cortez recently completed his Master in Public Administration at the Harvard Kennedy School, where he was an Ash Center Fellow and served as a research assistant at both the Bloomberg Center and the Institute of Politics. Before graduate school, he served as Senior Advisor and Deputy Chief of Staff in the Texas Congressional Delegation in Washington, D.C., where he helped shape legislative strategy for South Texas, built bipartisan consensus, and advanced policy on technology and economic development. He began his career as Defense and Veterans Liaison for U.S. Senator Kay Bailey Hutchison, later worked in corporate marketing, and subsequently co-founded a sports technology startup after business school. Josh has taught marketing and technology at the University of North Texas. He holds a Bachelor of Arts from Baylor University, an MBA from the University of Virginia's Darden School of Business, and an MPA from the Harvard Kennedy School of Government.

Shan Sherwan Hussein

Shan Sherwan is a Research Fellow at the Center for Public Leadership at the Harvard Kennedy School, where she focuses on negotiation research. She recently earned a Master in Public Administration from the Harvard Kennedy School. Shan is a specialist in gender-based violence and an experienced humanitarian leader. She began her career supporting Syrian refugees and internally displaced Iraqis during the war with ISIS. Over the years, she has contributed to humanitarian responses in Iraq, Ukraine, and Jordan, developing toolkits and resources to support women and girls affected by conflict.

Ala'a Kolkaila

Ala'a Kolkaila is a strategic advisor and economist with over a decade of experience in public policy, digital transformation, and economic development. She recently graduated from Harvard Kennedy School as a Mason Fellow and from MIT Sloan School of Management as a Sloan Fellow (MBA). In 2024, she was a Fellow at Harvard's Belfer Center, advancing research on emerging technology, policy, and global governance, and also served as a Millennium Leadership Fellow at the Atlantic Council. Ala'a has played a pivotal role in shaping long-term strategies for the Egyptian government and international organizations, contributing to initiatives that integrate technology and innovation into economic and environmental frameworks. Her work has influenced policies on mobile money, cryptocurrency, and cross-border aid, strengthening international cooperation and development agendas. Fluent in English and Arabic, Ala'a's work epitomizes her unwavering commitment to integrating innovative solutions with policy development to tackle complex global challenges head-on.

Rod Moshtagi

Rod is a Master in Public Policy candidate focused on addressing the societal risks posed by artificial intelligence, particularly in the areas of workplace displacement, disinformation, and the gradual disempowerment of human decision-making. His core mission is ensuring that the benefits of AI are not concentrated in the hands of a few, but are instead widely democratized for the public good. Rod was selected as a student leader in AI at the Berkman Klein Center for Internet & Society and worked as the workshop lead for the AI Safety Student Team at Harvard. Prior to his graduate studies, he worked as a consultant at PricewaterhouseCoopers, where he focused on AI and analytics initiatives, including projects related to responsible AI governance. Looking ahead, Rod is committed to advancing systems that empower individuals with AI, rather than replace them, and ensuring that technological progress enhances human agency and social cohesion.

Angie Sohn

Angie Sohn is a Master in Public Policy candidate and Belfer Young Leader at the Harvard Kennedy School concentrating in international and global affairs. She is broadly interested in diplomacy, emerging tech, and security issues with a focus on the Indo-Pacific region. A Thomas R. Pickering Fellow, Angie most recently worked in the U.S. Department of State's Bureau of East Asian and Pacific Affairs, in which she worked on the DPRK, ROK, and Mongolia country portfolios. She has previously served as a Fulbright ETA in Taiwan and graduated from Colby College with a BA in government and minor in East Asian studies.

Fabian Ulmer

Fabian is a Master in Public Policy (MPP) candidate and McCloy and Hans Weisser Fellow at Harvard Kennedy School, concentrating in International and Global Affairs (IGA). His research interests lie in the safe, ethical, and trustworthy development and use of Artificial Intelligence (AI). More concretely, he aims to foster

international cooperation in the global task to mitigate AI risks, align AI's innovative power with ethical and social responsibility, and ensure technology serves people through better governance. Prior to HKS, he contributed to the adoption of the European AI Act while serving in the Cabinet of Executive Vice-President Margrethe Vestager and supported the implementation of the legislation and its General-Purpose AI Code of Practice at the think tank The Future Society.

Liz Valdez

Liz Valdez is a Master in Public Policy candidate at the Harvard Kennedy School, concentrating in political and economic development. A Fulbright Scholar and John F. Kennedy Fellow, she has over four years of experience at the Mexican Ministry of Foreign Affairs, where she served as Officer for Political and Economic Affairs for North America and as Deputy Director for Migration Policy and Data Analysis. There she focused on Mexico's economic cooperation initiatives with the US and Canada, trade research including USMCA consultations, commercial diplomacy efforts, legal pathways and community integration for refugees and migrants, and promoting academic mobility. More recently, she worked with Harvard's Growth Lab on economic development in Morocco, exploring strategies for sustainable inclusive growth. Liz holds a BA in international relations from ITAM and has completed certifications on foreign policy, immigration law and international economics from UC San Diego, CIDE, University of Arizona and UNHCR.

Sam Weiss

Sam Weiss is a Master in Public Policy student and John F. Kennedy Fellow at the Harvard Kennedy School. During the summer of 2025, Sam served as the Foreign Policy and National Security Fellow for Senator Jacky Rosen (D-NV), where he focused on foreign assistance, counterterrorism, and the Middle East. Formerly Special Assistant to the Director at USAID's Office of Budget and Resource Management, Sam has advised senior leadership on strategic planning and budget execution for the multibillion-dollar humanitarian assistance and geostrategic infrastructure portfolios, shaping US responses to crises like those in Ukraine, Gaza, Haiti, and Sudan. He has managed global humanitarian programs, advanced digital development and cybersecurity efforts, and championed robust foreign aid efforts on Capitol Hill. Sam earned his BA in International Studies from the Ohio State University as a John Glenn Fellow, Buckeye Leadership Fellow, and the university's nominee for the Rhodes and Marshall Scholarships.

Acknowledgements

Table of Contents

Executive Summary

The rapid evolution of artificial intelligence (AI) technologies has transformed many sectors; conflict resolution and peacebuilding are no exception. This paper synthesizes real-world evidence and scholarly insights illustrating AI and digital technologies' capacity to improve traditional conflict resolution and peacebuilding methodologies while outlining their limitations.

The integration of AI into daily life has introduced novel ways to gauge public opinion, predict conflict escalation, and enhance inclusivity and effectiveness in mediation. AI's applications span from early warning systems and digital negotiation platforms to context analysis tools and deliberative simulations, all of which are reshaping the dynamics of contemporary conflict resolution and peace processes. Beyond improving efficiency, these tools can deepen stakeholder engagement and provide data-driven insights into underlying grievances, thus facilitating more adaptive and responsive peacebuilding strategies.

Alongside these advances, AI presents significant challenges, particularly across operational, ethical, and political dimensions. At the operational level, AI systems risk reinforcing historical biases, exacerbating existing power imbalances, and complicating the equitable distribution of technological benefits. Ethically, issues such as algorithmic biases, opaque decision-making processes, and the potential misuse of AI-generated content underscore the need for robust oversight mechanisms and comprehensive governance frameworks. This echoes across various political realms, where the adoption of AI in conflict resolution and peacebuilding raises concerns over sovereignty. This is exacerbated by unequal access to technological resources, and the instrumentalization of AI tools by powerful actors to advance their own interests, thereby undermining trust and inclusivity in peace processes.

This paper builds upon insights and discussions initiated by the AI+X+G study group at the Belfer Center for Science and International Affairs, Harvard Kennedy School, which served as a key catalyst for exploring the intersection of artificial intelligence and conflict resolution. It advocates for an integrated global response to these challenges, emphasizing the importance of robust ethical oversight, inclusive governance practices, and international regulatory frameworks that safeguard rights, promote transparency, and democratize access to AI technologies. Collaborative efforts among multilateral institutions, governments, academia, and civil society are essential to ensure that AI's transformative potential is harnessed responsibly, ultimately fostering inclusive and equitable outcomes in conflict resolution and peacebuilding.

Introduction and Overview

This paper uses the term "artificial intelligence" (AI) to refer primarily to systems that analyze and interpret data, make predictions, or support decision-making in ways that emulate aspects of human cognition. This definition includes machine-learning-based AI, large language models (LLMs), and language reasoning models (LRMs) as well, given their increasing relevance in processing complex information and facilitating communication within conflict resolution contexts. While there are some references to digital tools such as translation software or blockchain technologies, the paper distinguishes these from AI systems unless otherwise noted.

Peacebuilding refers to comprehensive, long-term efforts to prevent the outbreak or recurrence of violent conflict by strengthening national and local capacities for conflict management and resolution. It seeks to create the conditions for durable peace and development by addressing core issues that affect the functioning of society and the state.[1] Conflict resolution is the process of addressing the root causes of conflict and reaching mutually acceptable solutions through dialogue, negotiation, or mediation, with the aim of ending violence and laying the foundation for sustainable peace. Negotiation and mediation are closely linked; whereas negotiation involves parties directly, mediation brings in a neutral third party to support dialogue. The term "peacebuilding" is used as an overarching framework that includes, but is not limited to, conflict resolution. While conflict resolution typically refers to formal negotiations, mediation, or settlement of disputes, peacebuilding also encompasses longer-term efforts to address structural injustice, promote inclusion, and strengthen civic trust. In this paper, when "conflict resolution" appears, it refers to these more specific practices within the broader landscape of peacebuilding. Core principles for just, inclusive, and durable peace include addressing root causes of violent conflict, local capacity development, just and meaningful inclusion, accountability and transparency, and solidarity and social justice.[2]

This paper adopts a values-driven understanding of peacebuilding that goes beyond the cessation of conflict and encompasses inclusion, dignity, and agency. These principles inform the analysis of AI's role in supporting or undermining meaningful peacebuilding processes. Revolutionary progress in AI is transforming approaches to conflict resolution and peacebuilding. AI today is actively influencing the field with predictive models that forecast social unrest and generative tools that draft complex negotiation texts.[3, 4] This integration remains uneven, and is shaped by access, data gaps, and contested norms.[5] While the last wave of digital innovation changed how data is communicated and shared, today's AI signifies a substantial leap forward, offering the potential to improve mediation and peacebuilding in entirely new ways. However, alongside that promise comes risk. This paper explores how AI can both facilitate and complicate the future of conflict resolution and peace.

Conflicts today are complex and rapidly evolving, shaped by shifting alliances and deep-rooted socio-political tensions. Traditional negotiation frameworks, such as distributive, integrative, and principled approaches, face inherent limitations, including human cognitive biases, information processing constraints, and challenges in identifying underlying interests or objective criteria. AI offers significant improvements by enhancing data analysis and prediction, automating option generation, providing real-time strategic insights, and mitigating human biases.

Early experiments in AI for conflict resolution focused on rule-based logic and basic simulations.[6] However, significant progress has been made in recent years. Machine learning and natural language processing have opened the door to interpreting real-time dialogue, analyzing sentiment, and drawing insights from messy, unstructured conversations. These tools help to uncover some underlying issues, identify opportunities for consensus, and guide negotiators with fresh perspectives.

There is also growing momentum behind what some call "hybrid peacemaking intelligence" —a collaboration between human judgment and AI's computational power.[7] AI-driven tools can surface stakeholder positions, forecast negotiation paths, and even suggest strategy shifts. Properly designed, such hybrid models may help broaden participation and improve transparency in peace processes, raising the prospect that underrepresented voices and perspectives can be more effectively included. As explored later in this paper, the quality of peace is closely tied to the degree of inclusivity and transparency in its making.

AI operates within a wider ecosystem of technologies that carry inherent risks, including surveillance, data bias, and the exclusion of marginalized voices. These tools can entrench power just as easily as they challenge it. This dual nature of AI underscores the urgent need to introduce robust governance and ethical frameworks. In the context of conflict resolution and peacebuilding, with inadequate oversight, AI has the potential not only to amplify existing inequities but also to generate new forms of exclusion and vulnerability, thereby weakening trust, legitimacy, and inclusivity in peace processes. Addressing these requires more than technical fixes; it demands governance and accountability frameworks grounded in international human rights principles such as privacy, fairness, and due process.[8] In the absence of shared norms, states and institutions are moving quickly to define technical guidelines, often in ways that may reflect strategic interests or geopolitical priorities rather than participatory or rights-based deliberation. These dynamics are further shaped by the growing influence of private actors that control core AI infrastructures and training data. Therefore, for AI to be a responsible and effective tool in peacebuilding, its deployment must be grounded in diverse knowledge systems and tailored to the unique cultural, historical, and political realities of each conflict. This sensitivity is crucial to prevent the imposition of universal frameworks, which risk being counterproductive or detrimental when applied to complex local realities.

The sections that follow examine both the potential and the risks of AI in conflict resolution. They consider how predictive analytics are being used in conflict forecasting, how large language models are entering negotiation spaces, and why ethical frameworks and international coordination are essential. Drawing on real-world examples and academic research, the paper traces how AI is being deployed in practice, what is at stake, and how its use may shape the quality of peace.

Ultimately, this paper aims to advance a forward-looking yet balanced perspective. AI, used wisely, can make peacebuilding more effective and inclusive. But it must be guided by strong values of transparency, accountability, empathy, and justice. The mediation and conflict resolution community faces the challenge of integrating AI in a manner that enhances the human element at the core of our efforts.

Standards for Durable Peace

Before turning to how AI is entering peacebuilding practice, it is necessary to establish the standards against which its contributions should be assessed. The question is not only whether AI can speed up dialogue, broaden access, or improve prediction, but whether these innovations contribute to the kind of peace that endures or is sustained. Durable peace has long been understood, across scholarship and practice, to rest on principles of inclusion, accountability, justice, and transparency. These are not abstract aspirations. They are grounded in decades of comparative evidence on peace process outcomes, codified in international frameworks, and reinforced by practitioner experience that shows peace lasts when it is broadly owned, socially legitimate, and embedded in resilient institutions, and, conversely, that exclusionary or opaque processes tend to collapse or leave grievances unresolved.

It is therefore essential to clarify, before assessing technological applications, that AI should be judged by the same yardsticks that have guided peacebuilding more broadly. Efficiency, speed, or novelty cannot be the sole measures of success. The relevant test is whether AI-enabled processes uphold or exceed the standards that evidence shows produce durable, legitimate, and just peace. The following section reviews these established frameworks and findings.

Normative Frameworks

The UN has consistently affirmed that peace cannot be reduced to the cessation of armed violence. The 1992 Agenda for Peace[9] underscored that lasting peace depends on accountability, participation, justice, and the resilience of institutions. The Women, Peace, and Security (WPS) agenda, starting with UNSCR 1325 (2000)[10] and reinforced by subsequent resolutions, expanded this foundation and established that women's meaningful participation in peacebuilding is both a matter of rights and a determinant of effectiveness. Building on this mandate, the UN Guidance for Effective Mediation (2012) broadened the scope of inclusion to civil society actors, displaced persons, minorities, indigenous communities, victims' associations, and the incorporation of plural knowledge systems into decision-making.[11] The Gender and Inclusive Mediation Strategies guidance (2017) translates these principles into operational practice.[12]

The Secretary-General's 2023 A New Agenda for Peace[13] and the Pact for the Future[14] frame durable peace as inseparable from inclusive governance, justice, and rights-based accountability. Other recent commitments such as the Global Digital Compact and the Declaration on Future Generations[15] extend these principles further and embed digital cooperation into the agenda for peace and security. Within the 2030 Agenda, SDG 16 operationalizes these commitments by calling for inclusive societies, equal access to justice, and effective, accountable, and transparent institutions. [16] For this paper, these standards are important as they also define the benchmarks that AI-enabled tools should meet if they are to contribute to durable and legitimate outcomes.

Peace Process Evidence, Critique, and Design Trade-Offs

Comparative research and practice accounts converge that inclusion and transparency correlate with durability and legitimacy, but also show how these gains depend on how inclusion is designed and where it occurs. Three clusters of findings are salient for this paper: (i) empirical evidence linking broad participation to more stable agreements; (ii) design lessons on modalities and sequencing; and (iii) critiques showing the risk of symbolic participation, elite resistance, and opaque bargaining. These lessons are relevant for AI in peacebuilding, because the way systems are designed, i.e., what data they are trained on, how they process inputs, or how their outputs are used, will determine whether they broaden participation or replicate exclusion.

Cross-national studies and practitioner accounts both substantiate that meaningful inclusion lowers the risk of peace agreement failure. Rausch and Luu (2017) show that youth-inclusive dialogues in Nepal generated consensus and reduced violent demonstrations by 80 percent.[17] The creation of channels for nonviolent political expression connected grassroots actors with formal decision-makers, and produced joint recommendations that informed constitutional debates. They stress that inclusion worked not only by expanding voices at the table, but by institutionalizing follow-up mechanisms to monitor commitments, thereby reducing relapse into violence. O'Reilly, Ó Súilleabháin, and Paffenholz (2015) show that when women are included in a peace process, the peace agreement that results is 20 percent more likely to last at least two years, and 35 percent more likely to last for fifteen years.[18] This points to the need for feedback loops that connect participation

with real political outcomes. If AI is deployed to broaden participation, its effectiveness may be measured with the longevity of peace agreements, along with efficiency or scale.

Design research shows that *who* is included, *how* their input is structured, and *when* it occurs matters as much as whether inclusion happens at all. Schädel and Dudouet (2020) distinguish between horizontal inclusion (armed actors and potential spoilers) and vertical inclusion (social actors such as women's groups, youth, victims' associations). Both are essential. The former without the latter risks an elite cartel, while the latter without the former risks unenforceable agreements. Modalities range from direct representation and observer status to consultative forums, commissions, or public participation, but they only strengthen peace when linked to real influence. Sequencing is equally important: "Incremental inclusivity" works when early stages provide ownership and reduce spoiler incentives, while late or tokenistic gestures produce clauses that are merely symbolic.[19] The *Why It Worked* project, undertaken by the Center for Public Leadership's Negotiation and Conflict Resolution Collaboratory at HKS, reinforces these lessons. Drawing on comparative analysis of more than seventy years of protracted, asymmetric, and ethnonational conflicts, it finds that peace processes endure when they build legitimacy for all core actors, institutionalize trust-building measures or power-sharing, and address root causes and collective trauma. Agreements that denied legitimacy, sidelined grassroots actors, or ignored historical grievances tended to collapse, while those that combined broad participation with concrete trust-building proved more sustainable.[20]

Ghais (2022) reinforces this through comparative evidence. When civil society organizations were given observer status in Liberia's 2003 talks, they shaped reforms that constrained interim leaders and initiated truth-seeking. These provisions were later credited with stabilizing implementation. In contrast, the exclusion of civic actors in Chad's 2002 negotiations produced a thin amnesty-and-office-sharing deal that quickly unraveled. Including broad constituencies, from CSOs and victims' associations to major armed groups, correlated with reduced spoiling and more sustainable outcomes, provided their input is structurally tied to decision-making.[21] Palmiano Federer (2024) shows how formal gestures of inclusion in Myanmar backfired, as women and civil society actors lacked real influence over agenda or power distribution.[22] While the Bosnia/Dayton Accord ended mass violence, its opaque, elite-driven design entrenched rigid power-sharing arrangements and generated long-term institutional paralysis rather than renewal.[23] This points to the need for broad inclusion of actors in AI-assisted peacebuilding, not only in who is consulted but in how their perspectives are weighed, translated into outputs, and tied to actual decision-making authority.

Cuhadar (2020) cautions that while inclusion is now a near-universal aspiration in peace process design, it is frequently resisted. This resistance is not incidental but stems from in-group/out-group dynamics. Most resistance originates with political elites, though societal actors and rebel groups also engage in it, deploying tactics that may be implicit or elusive, direct or explicit, or coercive. The motivations for resistance also matter. Resistance arising from political competition or economic self-interest can sometimes be managed through incentive structures or safeguards. Resistance grounded in identity (such as when gender or ethnic inclusion is perceived as an existential threat) requires longer-term strategies that combine institutional design with ideological and cultural change.[24] Here, AI design too faces a parallel risk. If training datasets underrepresent marginalized groups or over-weight dominant perspectives, or if algorithms are optimized for consensus without accounting for exclusion, the result may be to institutionalize resistance by validating the same asymmetries that drive conflict instead of mitigating it.

Emerging Benchmarks for AI in Peacebuilding

If these lessons shape the substance of peace processes, they must also inform the frameworks that govern

the use of AI within them. There are some emerging benchmarks on the design and deployment of AI in conflict resolution and peacebuilding.

The UN locates artificial intelligence within a rights-centered, development-oriented frame and identified conditions under which AI can support peace. General Assembly resolution A/78/L.49 (2024) recognizes the potential of "safe, secure, and trustworthy" AI to advance all sustainable development goals, promote peace, reduce digital divides, and protect human rights, while calling on states and stakeholders to refrain from uses that cannot comply with international human rights law, and to uphold rights through the AI life cycle.[25] The Security Council's debate (S/PV.9381) in 2023 characterized AI as a double-edged capability, which can enhance monitoring, early warning, and peace operations, but also threatens peace through disinformation, surveillance, autonomy in weapons, and bias. There were repeated calls by delegations for transparency, human oversight, accountability, and inclusive governance that gives equal voice to developing countries.[26] First Committee deliberations in 2024 (A/C.1/79/L.43) extend this security track by emphasizing human-centric, accountable, safe, secure, and trustworthy approaches to AI with implications for international peace and security.[27]

Complementing these system-wide signals, the UN Department of Political and Peacebuilding Affairs Practice Note on malicious use of ICT in mediation specifies process safeguards, such as technology-sensitive conflict analysis, preparedness for information manipulation, secure communications, and practical checks, to protect the integrity of peace talks.[28] UN Women's Action Brief on AI for WPS adds that AI can either widen or narrow participation in peace efforts depending on design choices, reinforcing the case that women's meaningful participation and civil society inputs must be structured into AI-enabled practices and not treated as an afterthought.[29] UNESCO's Recommendations on the Ethics of Artificial Intelligence provides ethics-linked expectations such as transparency, accountability, and human oversight as complements to, and not substitutes for, rights-based obligations.[30] The Special Rapporteur on freedom of expression (A/73/348) sets out a legal architecture that translates directly into AI in peacebuilding; that states must respect, protect, and fulfill rights in AI design, procurement, and deployment; and that private companies bear responsibilities aligned with the UN Guiding Principles, including policy commitment, due diligence, prevention/mitigation, ongoing review, and remedy.[31]

Responsible innovation frameworks echo these principles. Boulanin, Brockmann, and Richards (2021) propose Responsible Research and Innovation (RRI) as an actionable approach for AI in international peace and security, and highlight co-creation with stakeholders, anticipatory governance, reflexivity, responsiveness, and transparency as key.[32] SIPRI and UNODA's Factsheet on Responsible AI for Peace and Security defines responsible innovation as anticipatory, inclusive, reflective, and preventive, and points to governance levers across the research and technology lifecycle.[33]

Taken together, these initiatives matter for two reasons. First, they confirm that inclusion, transparency, and accountability are not optional features but widely recognized standards in peacebuilding. Second, they show that the question is not whether AI should be used, but under what conditions and with what safeguards. With these benchmarks in mind, the next section turns to the current landscape of AI in conflict resolution and peacebuilding, demonstrating how tools for prediction, mediation, and simulation are being applied in practice, and what is at stake for the quality of peace.

Current Landscape

The growing integration of AI into peacebuilding is reshaping the tools available to negotiators and mediators, as well as the underlying architectures of decision-making, legitimacy, and control. From early warning

systems to deliberative simulators, AI is becoming a force that structures whose voices are heard, which futures are modelled, and how peace is reimagined. Some have welcomed these developments, while others push for institutional and ethical guardrails amid their hasty implementations. Nonetheless, technological advancements in AI have generated optimism about its technical potential contributions to peace and mediation[34]—though important questions about outcomes, legitimacy, and long-term impact remain.

Conflict Prediction and Early Warning Systems

Early warning systems (EWS) that aggregate and analyze social, economic, and political data have rapidly evolved, enabling the forecasting of potential violence several years in advance. These systems are designed not only to issue alerts but also to guide preventive diplomatic and humanitarian actions, enabling decision-makers and mediators to act in a more timely manner. By processing vast datasets in near real-time from sources like social media, news reports, and governmental data, AI-powered EWS can identify subtle signals and patterns that indicate social fragmentation or rising tensions, which might otherwise go unnoticed.

A prominent example is the **Violence & Impacts Early-Warning System (ViEWS),** a conflict prediction platform jointly developed by Uppsala University and the Peace Research Institute Oslo (PRIO).[35] ViEWS employs machine learning to forecast the risk and intensity of political violence up to thirty-six months in advance, including estimates of fatalities. For instance, in June 2025 it forecasted 194 fatalities in Syria and 465 in Ukraine for September 2025, and for February 2028, it predicted 452 and 117 fatalities, respectively.[36] These forecasts are probabilistic and rely on a range of structural and dynamic indicators, including political exclusion, past violence trends, and socioeconomic conditions, rather than scenario-specific assumptions (e.g., regime change).

The forward-looking nature of these predictions raises questions about their underlying basis. VIEWS' forecasts are not static; they are part of a dynamic system that produces rolling, monthly updates. This dynamic model is evident when comparing successive forecast runs. Between the May 2025 and June 2025 updates, the fatality prediction for September 2025 in Syria remained stable at 194, while the forecast for Ukraine in the same month dramatically increased from 252 to 465. The sharp revision in prediction for Ukraine reflects new data indicating an escalating risk. While these figures highlight the system's ability to identify regions at risk, they also underscore the key challenge of predicting the precise scale of violence. The forecasts for ongoing conflicts may seem low, indicating that while the system can accurately pinpoint where violence will occur, quantifying the exact human cost remains a significant hurdle.

Conflict Mediation and Peace Negotiations

Beyond prediction, AI is already being deployed during conflicts as a powerful companion for mediation processes, particularly those involving large and complex datasets. The United Nations Department of Political and Peacebuilding Affairs, led by innovators like Martin Wahlisch in its former Innovation Cell, has used **AI-driven digital dialogue platforms** in countries like Libya and Yemen.[37] These platforms process thousands of public responses simultaneously, allowing mediation teams to generate granular insights by systematically analyzing citizen sentiment across different languages, ethnicities, and genders. As Wahlisch notes, such tools effectively combine the scale of a survey with the detail of a focus group discussion. AI-facilitated digital dialogues have also enabled inclusive participation where traditional polling methods fall short, allowing mediators to bypass security-related barriers, rapidly capture dynamic opinions, and confer legitimacy on

peace negotiations.[38]

The development of specialized AI tools further illustrates this trend. One such tool, **Latent Issue Extraction,** employs advanced natural language processing (NLP) techniques, including topic modeling and BERT embeddings, to analyze extensive negotiation transcripts.[39] This analysis uncovers patterns of convergence and divergence in opinions among conflict parties, enabling the design of negotiation strategies that address previously unrecognized grievances.

Another tool, the **Party Distances Tool,** quantifies the ideological gap between negotiating parties to map out potential "Zones of Possible Agreement."[40] These tools enhance the mediator's capabilities by helping them manage information overload and inherent human biases through data-driven metrics. Additionally, visual tools, such as party activity graphs, analyze the dynamics of conversation, highlight dominant speakers, and identify shifts when negotiations begin to stall or accelerate.

These tools and innovations reflect a shift toward **Hybrid Intelligence Systems** that combine human oversight with machine analytics. Currently, these systems acknowledge that AI's strengths lie in processing large datasets and identifying patterns, while human mediators provide the essential cultural, contextual, and ethical perspectives necessary for making nuanced decisions. The interplay between AI-driven analysis and human judgment enables the generation of "reflexive knowledge" that can adapt to the evolving nature of conflicts. For instance, by continuously updating analysis based on real-time data, hybrid systems help maintain situational awareness even in protracted or rapidly shifting conflicts. More importantly, how that knowledge is interpreted also depends on who is involved in the process, and the perspectives and contexts they bring. Crucially, this collaborative approach also mitigates the risk of overreliance on machine outputs by ensuring humans remain "in the loop," adapting and interpreting AI-generated insights in a culturally and contextually appropriate manner. The current role of AI in conflict resolution is best understood as a partnership that enhances the capabilities of human mediators. The core of mediation remains deeply relational and political, requiring interpretation, trust-building, and attention to asymmetries. As one panelist at Harvard's 2024 AI Negotiation Summit noted, while frontline negotiators might tolerate AI hallucinations and biases, this cannot be a safe assumption for all users, reinforcing the need for expert human oversight.[41] However, this human-machine dynamic is not static. As AI technology rapidly evolves, we anticipate that future systems may develop more human-like capabilities, such as contextual understanding and nuanced discernment. This evolution also presents an intriguing possibility where AI systems could be developed to monitor human decision-making to flag potential biases our own context might inadvertently introduce. While human perspective is invaluable, it can also be a source of inherent bias; an AI that serves as an objective check could improve this symbiotic partnership.

Conflict Simulators and Strategic Modeling

Simulation tools, part of a longer lineage of tech-supported negotiation experiments, have emerged as a new frontier in conflict resolution. The **Ukraine-Russia Peace Agreement Simulator,** developed by CSIS using data from forty-five experts and 375 ceasefire agreements, enables users to test hypothetical outcomes across various themes, including territory, justice, and security.[42] The simulator was built using software from Scale AI and data from forty-five experts who participated in a strategy game called "Hetman's Shadow," which was enriched with media analysis and data from 375 past ceasefire agreements.[43] This gamified environment enables negotiators and analysts to explore plausible futures and decision trade-offs in real time.

Similarly, **Google DeepMind's Habermas Machine,** although not yet used in conflict zones, has demonstrated

that AI-generated deliberative statements can outperform those written by human negotiators in terms of clarity and acceptance.[44] In one experiment, 56 percent of participants preferred the AI-generated statements, finding them clearer and of higher quality. With further development, such tools could help facilitate consensus-building in complex disputes. While promising, such tools also raise the question of whose definition of consensus is being coded, and whether these platforms can accommodate plural, contested, or nonnegotiable claims that resist algorithmic resolution.[45] For AI to benefit peacebuilding globally, particularly in regions with infrastructure limitations, targeted investment in public education and digital infrastructure is essential to ensure not just participation but meaningful voice in shaping the aims and boundaries of these tools.

Stakeholders in AI and Peacebuilding

The use of artificial intelligence in conflict resolution is cross-sectoral, demanding active collaboration among different actors. These stakeholders, ranging from multilateral institutions and national governments to private companies and grassroots organizations, may not always share common incentives. Understanding the power dynamics, capabilities, and limitations of each group is essential to designing ethical, inclusive, and effective AI systems for peacebuilding.

Below, we outline five primary stakeholder categories, each playing a distinct role in shaping the development, governance, and deployment of AI in conflict settings.

1. Multilateral Institutions

These institutions play a crucial role in establishing global norms, promoting cooperation, and advocating for the ethical use of AI in conflict resolution and peacebuilding. However, funding politics and limited technical capacity often constrain their influence, making them structurally dependent on the technical pipelines controlled by private firms and powerful states.

- **United Nations (UN):** The UN's engagement with AI in peacebuilding spans different dimensions. The UN Department of Political and Peacebuilding Affairs (UN DPPA) has used AI to conduct public sentiment analysis and facilitate multilingual "digital dialogues" in conflict-affected communities such as in Guatemala, Yemen, and Libya.[46] Training initiatives like the UN System Staff College's digital literacy programs further support internal capacity-building.[47] At the policy level, the UN Institute for Disarmament Research (UNIDIR) launched a series of global conferences and dialogues on AI, Security, and Ethics to foster international consensus on responsible AI development in the security domain.[48] These forums aim to build a shared understanding of risks and opportunities, informing the ongoing UN-led discussions on international security and AI governance. On the research and field practice front, collaborations with other organizations such as Interpeace have yielded AI-informed disinformation mapping and digital platform design in Africa.[49] At the normative and governance strand, the UN system has spearheaded efforts such as the 2024 UNGA Resolution on "Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development."[50] In addition, the Inter-Agency Working Group on AI, co-led by the ITU and UNESCO, remains a key internal knowledge-sharing forum.[51]

- **Regional and Member-Based Multilaterals:** The African Union's Peace and Security Council[52] and

intergovernmental partnerships like the OECD's Global Partnership on AI (GPAI)[53] promote regional cooperation and introduce perspectives grounded in specific regional priorities. The AU also developed the Continental Artificial Intelligence Strategy in 2024 to "leverage AI in a manner that promotes social and economic development and preserves its peace and security." In Southeast Asia, ASEAN has advanced regional frameworks through the ASEAN Guide on Governance and Ethics[54] and the ASEAN 2025 Digital Masterplan[55], along with the ASEAN-US Leaders' Statement on Promoting Safe, Secure, and Trustworthy Artificial Intelligence[56], which collectively set standards for responsible AI adoption and digital cooperation in the region. These initiatives highlight that AI in peacebuilding must remain context-sensitive rather than based on one-size-fits-all models.

2. National Institutions

Nations are expanding investments in AI research, but their roles vary widely, from regulation to prioritizing military applications. Bureaucratic silos can challenge effective integration, and in some cases, AI adoption may be driven more by modernization than conflict resolution needs, including applications in surveillance and predictive operations that may lack sufficient civilian oversight.

- **Ministries:** Switzerland's Federal Department of Foreign Affairs (FDFA) established the Geneva Science and Diplomacy Anticipator (GESDA) to bridge government, science, and the private sector.[57] In February 2025, South Korea's MOFA launched an "AI Diplomacy Division" to address AI's impact on global affairs, including security and economic growth.[58] Governments, however, differ in their objectives for AI adoption. While some emphasize applications for peacebuilding, others focus on areas such as national security, surveillance, or geopolitical advantage. In many cases, adoption is shaped by broader modernization agendas rather than by explicit peacebuilding goals.

- **AI Safety Institutes (AISI):** Following the AI Seoul Summit 2024, signatory states pledged to establish a network of AISIs to coordinate cross-border research and develop safety standards.[59] France's creation of the National Institute for AI Evaluation and Security (INESIA), announced at the 2025 AI Action Summit, marked a significant step toward public sector leadership in AI risk governance.[60] While in the early stages, these institutes have begun drafting baseline protocols, although the effectiveness of their enforcement remains to be seen.

3. Civil Society and NGOs

Civil society actors play a dual role: They implement peacebuilding tools and serve as watchdogs of AI ethics and harms. However, they are too often excluded from the initial design and procurement phases of AI systems. This exclusion is not merely an oversight but stems from closed, proprietary development cycles or government procurement processes that lack public consultation.

The most important role for civil society is not necessarily in the technical design (i.e., coding the algorithm) but in shaping the system's requirements, constraints, and ethical guardrails. They possess the crucial ground-truth knowledge to inform what a tool should do, what it must not do, and how its impact should be measured. Yet they are frequently brought in only after deployment to mitigate damage or manage backlash.

Finally, and most critically, there are the stakeholders who often lack a formal voice yet bear the greatest risks: the conflict-affected and marginalized communities themselves. Although they are frequently

categorized as beneficiaries rather than active participants, these groups are essential stakeholders whose safety and well-being represent the ultimate measure of success. They face increased risks from algorithmic bias and exclusion from decision-making processes, making their active participation and protection imperative for any legitimate and sustainable peacebuilding effort.

- **Humanitarian and Mediation Organizations:** Groups like Conciliation Resources and the International Committee of the Red Cross (ICRC) are developing frameworks for ethical AI use in humanitarian and conflict settings.[61] The ICRC's work, for instance, focuses on ensuring AI tools used in humanitarian action comply with international humanitarian law, particularly the principles of distinction, proportionality, and precaution. The framework is in an active development and piloting phase, representing an effort to establish clear operational rules before technologies become widespread. Since April 2023, the CMI-Martti Ahtisaari Peace Foundation has used AI for broad-based consultations and digital dialogues with stakeholders in Sudan.[62]

- **AI Safety Organizations:** Organizations such as AI for Peace and the Future Society promote the ethical deployment of AI, transparency, and multi-stakeholder engagement.[63] The Netherlands AI Coalition has a dedicated working group on security, peace, and justice.[64] These groups often act as intermediaries, connecting communities, governments, and developers.

4. Think Tanks and Academia

Academic institutions contribute critical insights into the technical development and ethical implications of AI in conflict settings.

- **Think Tanks:** The Center for Strategic and International Studies (CSIS) and the Center for a New American Security (CNAS) explore AI models for peace negotiations.[65] Security-focused institutions like the RAND Corporation also contribute significantly, often focusing on AI's role in national security, deterrence, and military strategy, which shapes the broader geopolitical context of its use.[66]

- **Academic Institutions:** Institutions like NYU's Peace and Research Education Program (PREP)[67] and Harvard Kennedy School's Belfer Center[68] facilitate scholarship at the intersection of AI, ethics, and international relations. Uppsala University and the Peace Research Institute Oslo (PRIO), through the ViEWS project, have advanced machine-learning models for predicting political violence.

5. Private Sector and Big Tech

Private technology companies are the primary developers of advanced AI systems, given their concentration of R&D capacity, computing power, and proprietary data. These firms increasingly participate in public-private partnerships, though their commercial incentives do not always align with public policy objectives. Market-driven priorities such as scale and efficiency can at times overshadow considerations of fairness, accountability, or cultural adaptation, which are important for sustainable peace.

- **Big Tech Companies:** Companies like Microsoft, Google, Meta, and Amazon have developed AI tools with implications for conflict resolution. Google's Jigsaw and DeepMind's experimental Habermas Machine aim to support deliberation and combat harmful speech. Microsoft's AI for Good initiative has supported disaster response and humanitarian mapping tools in partnership with the UN.[69]

- **AI Startups:** Emerging firms are designing specialized tools for early warning systems, negotiation

analytics, and conflict monitoring. While more agile than large corporations, these startups often rely on public or philanthropic funding and may lack the scale to achieve widespread adoption.

Opportunities and AI's Impact

Artificial intelligence offers considerable opportunities to strengthen conflict prevention, enhance mediation, and improve inclusivity and responsiveness in peace processes. These opportunities arise not only from AI's technical capacity to process large-scale data but also from its potential to augment human judgment, foster participatory frameworks, and support ethical, data-driven interventions in complex environments.

1. Predictive Intelligence and Governance for Conflict Prevention

AI-powered early warning systems analyze vast, heterogeneous datasets—including social media feeds, economic indicators, historical conflict trends, and satellite imagery—to detect weak signals that may precede violent conflict or political instability.[70] These insights can enable timely, targeted diplomatic or humanitarian interventions. AI also allows for granular geographic and temporal forecasting in regions previously underserved by conventional monitoring efforts.

The predictive power of these systems depends not only on data quality but also on interpretive context, especially regarding who is warned, about what, and with what consequences. These decisions may reflect institutional or geopolitical priorities rather than local needs, creating a potential tension between supra-national risk assessments and community-level realities. The deeper conflict, therefore, lies not just in geography or scale but in asymmetries of data access, technological capacity, and epistemic authority. While information gaps have always been a feature of international relations, AI supercharges this disparity. It enables technologically advanced actors to process data at an unprecedented scale, widening the information gap and creating a more profound imbalance than ever before.

Hence, to be effective for conflict prevention, predictive systems can embed inclusive governance frameworks which can not only ensure high-quality data but also integrate local perspectives into analysis, clarify decision-making protocols, and align preventive action with community priorities. Done well, AI-enabled early warning can help address structural drivers of conflict, such as economic exclusion, political marginalization, or environmental degradation, before they escalate, while also guiding resource allocation, inclusive dialogue, and counter-disinformation efforts in fragile settings.

2. Augmented Mediation Through Hybrid Intelligence

AI can powerfully augment the capabilities of human negotiators, helping them process information, identify patterns, and generate insights. Hybrid intelligence systems, which combine human oversight with machine analytics, are a particularly significant contribution.[71]

Such hybrid intelligence systems facilitate the generation of "reflexive knowledge"—a dynamic feedback loop where AI provides data-driven insights, humans interpret them based on context and ethics, and this refined understanding is then fed back to guide further analysis. For example, real-time analysis of sentiment trends or ideological divergence can inform tactical adjustments during negotiations. In long-running or personnel-fragmented processes, AI can help maintain situational awareness by tracking

changes in stakeholder positions and dialogue tone over time.

These systems also offer an opportunity to strengthen human judgment. When designed intentionally, they can help negotiators reflect on implicit biases, interrogate dominant narratives, and make more deliberate choices about whose voices and histories are included in peace processes. Leveraging AI to highlight underrepresented perspectives, elevate local knowledge systems, and counter algorithmic drift can make conflict resolution efforts more accountable and adaptive. Far from replacing the human element, AI can amplify it by surfacing insights that might otherwise be overlooked, challenging dominant framings, and expanding the ethical range of mediation practice in complex settings.

3. Increased Inclusivity

AI-powered tools offer concrete solutions to barriers that have historically hindered participation in peace processes. Rather than merely providing access, AI can actively foster inclusion. For instance, advanced, AI-driven machine translation can overcome language barriers in real-time, supporting dialects and colloquialisms that standard tools miss. Furthermore, AI can analyze qualitative data from thousands of participants in digital dialogues, using sentiment and topic analysis to ensure that the specific concerns of women, youth, and other marginalized groups are not just heard but systematically understood and presented to negotiators.

During the 2020–21 Libyan peace talks, the UN Support Mission in Libya (UNSMIL) convened virtual meetings to bring the voices of women and youth into a process dominated by seventy-five unelected representatives.[72] While virtual meetings provided access, AI offers the next step: the ability to analyze the rich input from such sessions at scale, ensuring the findings meaningfully inform the official political dialogue. However, to be genuinely inclusive, these tools must give communities control over how their data is used and represented, moving beyond token participation to genuine digital autonomy.

4. Improved Efficiency and Effectiveness

AI can significantly enhance the efficiency of peace processes by streamlining data analysis, facilitating communication, and modeling potential outcomes. But increased efficiency is not always value-neutral. Gains in speed or scale may come at the expense of deliberation, inclusion, or trust-building, particularly when complex political or cultural judgments are reduced to algorithmic outputs. Thus, understanding which forms of efficiency are pursued, and whose goals they serve, is crucial. Simulation platforms, like the CSIS Ukraine-Russia Peace Agreement Simulator, serve as a "sandbox" for refining strategies. AI can also enhance institutional memory by archiving negotiation histories, thereby improving continuity in long-running processes. This enhanced efficiency extends to post-conflict phases, where AI-powered systems using satellite and drone imagery can monitor and verify ceasefire compliance, building confidence through transparent monitoring. AI can also make peacebuilding more effective by:

- **Optimizing Resource Allocation:** AI models can analyze real-time needs assessments to help guide the strategic deployment of humanitarian aid and peacekeeping assets to where they are most needed.

- **Countering Misinformation:** Advanced algorithms can detect and flag deepfakes and disinformation campaigns in real-time, protecting the integrity of public discourse during sensitive negotiations.

- **Expanding Access to Support:** In low-capacity contexts, user-friendly AI tools can empower a more diverse range of local actors to contribute meaningfully to peacebuilding without relying on large technical teams.

- **Mitigating Bias:** When carefully designed, AI systems can help reduce certain forms of human bias in analysis and decision-making. For example, AI can identify patterns that might otherwise be missed in more intuitive or manual processes. However, AI systems' effectiveness depends on the data they are trained on and the assumptions built into their design.

Challenges

While AI offers tools with potential relevance for peacebuilding, its use in conflict resolution presents significant challenges. These challenges span cultural, technical, ethical, and operational dimensions.

1. Bias and Representational Limitations

One major challenge for AI in conflict settings is the representativeness of the data used to train or fine-tune models. Data from conflict zones is often incomplete, contested, or shaped by misinformation, which can distort how models interpret on-the-ground realities.

- **Cultural Misunderstanding:** AI systems trained primarily on general or global datasets may misrepresent culturally specific values, beliefs, or historically rooted grievances.[73] In peace processes, this can result in outputs that flatten plural worldviews into reductive compromises, worsening mistrust or undermining legitimacy for marginalized parties or minority groups.[74]

- **Reinforcing Inequity:** When training data disproportionately represents one group's perspectives, AI outputs may naturally favor that group. Recognizing this potential risk, the UN's use of AI to analyze the Israel-Palestine conflict did not rely on the model to deliver definitive answers, but the tool explored potential solutions and alternative options—acknowledging the importance of human judgment in navigating biased or unreliable information.[75] The outcomes of such models are constrained by the accuracy of the model and the weights it allocates to the different variables as much as its ability to distinguish between credible information and rumors or propaganda. Failure to do so risks reinforcing existing biases rather than offering accurate conflict analysis or viable pathways for resolution.

2. Digital Divides and Power Imbalances

Effective AI requires massive amounts of data and robust digital infrastructure, which are not universally available. This creates a fundamental gap.

- **Infrastructure Disparity and Data Divide:** AI systems require high-quality, timely data and reliable digital infrastructure, but these remain unevenly distributed. More than 2.7 billion people, mostly in the Global South, remain offline, leaving them underrepresented in AI-driven peacebuilding models due to the limited digital information available for reference.[76] This risks excluding entire regions from monitoring and mediation tools while reinforcing geopolitical inequities in whose conflicts receive

predictive attention or intervention.[77] In such cases, where datasets lack critical variables, AI models may not only produce biased outputs but also inaccurate recommendations, for example, overlooking signs of escalation in conflict-prone areas and thereby reducing the effectiveness of preventive or resolution efforts.

- **Epistemic Power and System Design:** Concentrations of expertise in peacebuilding, whether in AI systems or in human teams, risk embedding the norms and priorities of a narrow set of actors. With AI, this challenge is amplified, as design choices made largely in high-income countries can be codified into systems and applied globally, without adequate global input. Ensuring that AI incorporates plural sources of expertise is therefore essential to avoid reproducing bias and to sustain legitimacy in conflict settings.

- **Lack of Accountability and Redress:** Even when individuals participate digitally, opaque algorithms can mediate how their data is interpreted and acted upon, often without their knowledge or consent.[78] Without mechanisms for appeal, correction, or collective oversight, these systems may reproduce the same harms that they claim to fix—only faster, and with less transparency. This raises concerns about procedural fairness, dignity, and the governance of algorithmic authority in post-conflict settings.

3. Misinformation, Ambiguity, and Dual-Use Risks

AI tools can be both corrupted by bad information and used to create it, making them a double-edged sword in conflict. In conflict situations, this can lead to increased confusion, distrust, and a higher likelihood of escalation.

- **Misinformation:** AI systems are not only vulnerable to biased inputs but are increasingly being used to operationalize disinformation. In conflict settings, actors with access to information infrastructure can exploit model ambiguity to elevate specific narratives, frame adversaries, or undermine peace efforts. Additionally, the speed and ease with which AI can generate manipulated images, videos, and audio has significantly increased the velocity and volume of misinformation online.[79] Without safeguards, AI tools risk becoming amplifiers of strategic falsehoods.

- **Algorithmic Ambiguity:** Malicious actors can use AI to create highly realistic "deepfakes" and other synthetic content that blurs the line between truth and fabrication. This "corrosive ambiguity" erodes the trust needed for dialogue. The opacity of many AI models (often referred to as "black box" systems) poses risks for peace processes that depend on transparency and mutual accountability.[80] When creators of the tech cannot fully explain how the AI tool reached a specific conclusion, nor can adversaries audit how a recommendation was generated, AI outputs may be viewed with suspicion rather than propel resolution.

- **Dual-Use Risks:** Many AI systems are dual-use, meaning that those developed for peacebuilding may be entangled with military and security infrastructures.[81] Predictive tools used to assess conflict risk can also be deployed for surveillance, population control, or autonomous targeting. For example, satellite imagery that is used to monitor ceasefire compliance may be rerouted to inform drone strikes. Language models that are trained to understand grievances may be used to manipulate public sentiment or target opposition. These tools' repurposing can blur the line between peacekeeping and coercion.

4. Privacy and Surveillance Risks

In conflict settings, the collection of personal data to power AI systems, such as geolocation, social media activity, and satellite imagery, raises privacy and safety concerns. While such data can help with early warning and targeted responses, it also introduces the risk of surveillance, data misuse, and loss of agency for affected communities. Without enforceable safeguards, these tools risk compromising safety, eroding trust, and undermining the legitimacy of peacebuilding efforts.

- **Surveillance:** In places like Darfur, the United Nations Early Warning System was deployed to monitor the conflict. Analyzing satellite imagery has helped identify displaced people and target humanitarian aid. However, this level of monitoring creates a risk of "surveillance humanitarianism," where personal data is collected without proper safeguards.[82] This information could be misused by political actors or stolen in a data breach by adversaries.

- **Free Speech vs. Safety:** Online hate speech has been used to incite violence, intimidate vulnerable groups, and undermine peace efforts, especially in conflict zones. At the same time, digital spaces serve as critical platforms for political dissent and community organizing. AI systems intended to combat online hate speech must therefore maintain caution and ensure that safeguards against incitement do not suppress legitimate expression. Forums like the 2025 AI Action Summit have called for accountability mechanisms that protect both safety and civic participation in high-risk environments.[83]

5. Overreliance on Technology and Diminished Human Agency

A recurring concern is "AI solutionism," the belief that technology can resolve deeply human problems simply by analyzing data at scale. As Jeff Seul and others caution, this mindset encourages blind trust in models, treating algorithmic outputs as authoritative rather than advisory.[84] [85] In peacebuilding, this can displace the relational and deliberative processes that confer legitimacy on agreements.

There is a significant risk that negotiators may begin to over-rely on AI, treating algorithmic recommendations as authoritative and allowing automated outputs to shape the sequence, framing, or perceived legitimacy of negotiations. Over time, this can erode trust, weaken accountability, and displace the deliberative processes essential to sustainable peacebuilding.

- **The Illusion of Control[86]:** AI systems offer speed and scale but not understanding. While they can detect statistical patterns, they cannot grasp the lived experience of historical trauma, moral injury, or cultural red lines.[87] In fields like humanitarian assistance or refugee adjudication, deferring to algorithmic outputs without safeguards risks depersonalizing decisions and eroding accountability.[88] The risk is not that AI is inferior; these tools should support—not replace—the relational and adaptive work at the heart of peacebuilding.

- **Loss of Adaptability:** Peace negotiations often require flexibility and context-specific judgement, especially when talks reach an impasse. Overreliance on AI-generated scripts or recommendations can dull these capacities, which makes it harder for negotiators to respond to evolving dynamics or shifts in tone, trust, or alignment among parties. When mediation is treated as a technical exercise, rather than a human process, it risks prioritizing procedure over understanding.[89]

6. Governance, Implementation, and Sustainability Challenges

The governance of AI in peacebuilding is underdeveloped relative to its rapid adoption. While AI tools are increasingly deployed in conflict settings, global regulatory frameworks are fragmented, with most efforts resting on voluntary principles or national guidelines.[90] In peacebuilding contexts, this imbalance raises the risk that AI systems reproduce existing hierarchies instead of supporting inclusive, locally grounded responses.

- **Weak Rules and Oversight:** Unlike human negotiators, for whom norms of international law, mediation practice, and institutional mandates provide some guardrails, there are few enforceable rules governing AI in conflict resolution and peacebuilding. Current frameworks rely on nonbinding principles or fragmented national standards.[91] This governance vacuum increases the likelihood that AI systems are deployed in ways that are opaque, misaligned with human rights obligations, or disconnected from peacebuilding values such as inclusion and transparency.

- **Implementation Gaps:** Peacebuilding organizations often lack the technical infrastructure and institutional capacity required to effectively integrate AI systems. Even when tools are available, they are frequently developed by actors in the Global North, with limited input from those working in or affected by conflict. This disconnect can result in tools that are poorly aligned with local needs or cultural realities. Without stronger mechanisms for co-development, these gaps risk undermining the credibility, uptake, and impact of AI in peacebuilding.

- **Security Risks:** AI systems deployed in conflict zones are vulnerable to cyberattacks or manipulation. If these systems are compromised, they could provide misleading information or be used to harm rather than help.

- **Sustainability Issues:** AI systems must be continuously updated to remain useful in rapidly evolving conflict environments. However, keeping them running requires money, skilled staff, and a long-term commitment. In resource-limited settings, this is often not realistic. Without sustained investment, even well-designed systems can become outdated over time.

Addressing these governance, implementation, and sustainability issues requires establishing robust ethical guardrails, ensuring human oversight, and designing AI systems in collaboration with the communities they are intended to serve.

Policy Recommendations

To responsibly integrate artificial intelligence into conflict resolution efforts, we propose focusing on three core pillars: establishing robust governance and ethical standards, building inclusive infrastructure with equitable funding, and ensuring access and decision-making power are shared with communities directly affected by conflict. The recommendations reflect both near-term feasibility and long-term strategic goals:

1. Governance and Ethical Standards

I. Create Multilateral Task Forces on AI and Conflict Resolution

Led by: Multilateral bodies

Multilateral bodies such as the United Nations, OECD, African Union, and European Union should form dedicated task forces. These task forces must include representatives from conflict-affected regions and civil society to develop evolving ethical guidelines, data management standards, and rules for procuring commercial AI tools.

II. Enforce Independent Audits and Algorithmic Accountability

Led by: National governments, tech firms, multilateral agencies

National governments and multilateral agencies should establish task forces on AI and conflict resolution that set standards for transparency for private entities developing AI tools. Rather than acting as a traditional regulator, these task forces would leverage the power of procurement, embedding these standards into all contracts for AI tools used in peacebuilding. These procurement requirements would mandate independent, third-party audits to assess for bias, transparency, and potential misuse, with public disclosure of the findings. This allows the task forces to enforce the standards effectively by having the authority to void contracts if private entities fail to meet them.

III. Establish Minimum Safety Thresholds for Deployment

Led by: States and multilateral bodies

States and multilateral bodies should adopt binding safety standards for any AI tool used in conflict resolution. AI systems used in fragile or conflict-affected environments should undergo legal and ethical review based on international humanitarian and human rights law. This review ensures compliance with fundamental principles like distinction and proportionality, adapting the standards as the technology evolves.

In addition, such AI tools should be subjected to adversarial threat modelling and "red team" simulations, testing for malicious misuse by criminal actors, rogue states, or internal sabotage. This includes simulating how AI tools could be co-opted for surveillance, manipulation, or coercion. Safeguards like human override mechanisms, explainability requirements, and testing under contested conditions should be embedded in the design and procurement process, rather than being retrofitted after harm occurs.

2. Infrastructure and Funding

IV. Create a UN AI and Conflict Resolution Fund

Led by: The United Nations

The United Nations Department of Political and Peacebuilding Affairs should establish and fund an AI and Conflict Resolution Fund. In addition to advancing research on the use of AI in conflict resolution, it should prioritize projects focused on data quality and multilingual access to prevent failures in linguistically diverse conflict zones, as well as to avoid misrepresenting local nuances and excluding entire communities. Data collection must follow the ethical and integrity-driven guidelines established by the relevant task force as described above.

There should be special provisions for underserved areas, such as the Global South, where access to such technologies is limited.

V. Develop Global Procurement Guidelines for AI in Peacebuilding

Led by: Governments, international NGOs, UN agencies

Conflict-affected governments and multilateral bodies should adopt peace tech procurement standards requiring open model audits, dual-use disclosures, ethics vetting, and interoperability with human-in-the-loop frameworks. These should be modeled on arms control precedents.

VI. Support Anti-Disinformation Infrastructure in Peace Processes

Led by: Civil society, tech platforms, monitoring bodies

Disinformation, especially via deepfakes and AI-generated propaganda, can derail peacebuilding. Support must go to tools that detect and counter these threats in real time, focusing on empowering local organizations to lead response efforts. International mediators and donors should invest in tools and institutions that detect, archive, and counter synthetic media and deepfakes in conflict zones. This includes partnerships with fact-checking coalitions, metadata authentication protocols, and civic media training.

3. Access and Inclusion

VII. Invest in Digital Infrastructure and AI Literacy

Led by: National governments, development agencies, and education ministries

National and local governments should invest in digital infrastructure and public education around AI and conflict resolution. Given the need for AI literacy among negotiators, communities, and civil society groups to ensure equitable and informed participation in conflict resolution initiatives, these investments are critical to the success of AI tools in resolving future conflicts. AI literacy must go beyond technical training to include media literacy and disinformation awareness. Peacebuilders, journalists, and communities need tools to detect and critically evaluate synthetic media. Without

digital inclusion, AI could deepen divides rather than close them.

VIII. Ensure Participatory AI Design and Data Governance in Conflict Zones

Led by: Tech developers, international NGOs, peacebuilders

Multilateral agencies and donors should require participatory co-design practices in peace tech development. This includes embedding local actors in model training, establishing community data stewardship protocols, and institutionalizing pathways for redress of epistemic harm (misrepresentation, erasure, or misuse of local knowledge).

Conclusion

Conflict resolution and peacebuilding, whether led by people or assisted by AI, must be human-centered, ethically governed, and accountable to those most affected by conflict. The real measure is not the tool itself but whether the process safeguards the dignity and agency of those most affected by conflict. Integrating artificial intelligence into conflict resolution and peacebuilding is neither inherently liberatory nor inherently harmful. Its value depends on how it is designed, deployed, and governed, and by whom. Peace is not merely the absence of violence, but a condition of inclusion, dignity, and agency. AI systems, in their applications for conflict resolution and peacebuilding, must be assessed against these benchmarks, and not just technical benchmarks of accuracy, speed, or scale. At the same time, this paper has underscored the range of opportunities AI offers—from strengthening early warning systems to enhancing participation and strategic insight in peace processes. When developed and deployed with care, these tools can expand the reach and responsiveness of peacebuilding efforts.

The issue at hand is not whether we must innovate, but that we must make informed choices regarding how we innovate. The intention is not to slow progress or block innovation. The aims to take seriously the stakes of getting it wrong: the erosion of human judgment, the outsourcing of power to unaccountable systems, and the normalization of surveillance in spaces where trust is already fragile. Addressing these risks requires: enforceable safeguards, transparent procurement, independent oversight, and inclusive design processes that elevate local knowledge. Used well, AI can support human negotiators, expand access, and help manage complexity. But peace processes demand more than just optimization; they demand human participation, discernment, and care. If AI is to support peace, it must remain in service to these principles—not as a substitute for them.

Endnotes

1   From Terminology. (n.d.). United Nations Peacekeeping. https://peacekeeping.un.org/en/terminology

2   From Stimson Center. (2023, September 1). III. Peacebuilding Innovations for Global and Regional Problem-Solving from Future of International Cooperation Report 2023: Building Shared Futures: Innovating Governance for Global and Regional Problem-Solving. https://www.jstor.org/stable/resrep53183.9?searchText=why+are+inclusive+peace+processes+better&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3Dwhy%2Bare%2Binclusive%2Bpeace%2Bprocesses%2Bbetter%26so%3Drel%26sd%3D2015%26ed%3D2025&ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&refreqid=fastly-default%3Af418b8976b077b66669c4837ccadbfbb&seq=2

3   Sanguila, J.-A. P. (2024, October 9). Harnessing AI in locally-led peacebuilding efforts | Diplomatic Courier. https://www.diplomaticcourier.com/posts/harnessing-ai-in-locally-led-peacebuilding-efforts

4   Albrecht, E., Fournier-Tombs, E., & Brubaker, R. (2024). Disinformation and Peacebuilding in Sub-Saharan Africa: Security Implications of AI-Altered Information Environments. United Nations University. https://unu.edu/sites/default/files/2024-11/disinformation%20peacebuilding%20subsaharan%20africa.pdf

5   Ashby, H. (2023, December 6). A Role for AI in Peacebuilding. United States Institute of Peace. https://www.usip.org/publications/2023/12/role-ai-peacebuilding

6   Zeleznikow, J. (2021). Using Artificial Intelligence to provide Intelligent Dispute Resolution Support. Group Decision and Negotiation, 30(4), 789–812. https://doi.org/10.1007/s10726-021-09734-1

7   Hirblinger, Andreas T. 'When Mediators Need Machines (and Vice Versa): Towards a Research Agenda on Hybrid Peacemaking Intelligence'. International Negotiation 28, no. 1 (25 January 2022): 94–125. https://doi.org/10.1163/15718069-bja10050.

8   Pauwels, E. "Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention: Opportunities and Challenges for the International Community." Global Center on Cooperative Security, 2020. http://www.jstor.org/stable/resrep27551.

9   United Nations Executive Office of the Secretary-General (EOSG). (1992). *A/47/277 An Agenda for Peace | Preventive Diplomacy and Peacekeeping (Report of the Secretary-General pursuant to the statement adopted by the Summit Meeting of the Security Council on 21 January 1992)*. United Nations General Assembly Security Council. https://docs.un.org/en/A/47/277

10   UN Security Council. (2000). Security Council Resolution 1325. https://documents.un.org/doc/undoc/gen/n00/720/18/pdf/n0072018.pdf

11   UN Mediation Support Unit. (2012). Guidance for Effective Mediation. https://peacemaker.un.org/sites/default/files/document/files/2022/08/guidanceeffectivemediationundpa2012english0.pdf

12   UN Department of Political Affairs (DPA). (2017). Guidance on Gender and Inclusive Mediation Strategies. http://peacemaker.un.org/sites/peacemaker.un.org/files/1.%20English%20-GIMS_0.pdf

13   United Nations Executive Office of the Secretary-General (EOSG). (2023). A New Agenda for Peace | Our Common Agenda Policy Brief 9 (UN Executive Office of the Secretary-General (EOSG) Policy Briefs

and Papers). https://peacemaker.un.org/sites/default/files/document/files/2024/08/our-common-agenda-policy-brief-new-agenda-peace-en.pdf

14   UNGA Resolution A/Res/79/1 The Pact for the Future | 79th Session of the UNGA (2024). https://docs.un.org/en/A/RES/79/1

15   Pact for the Future, Global Digital Compact, and Declaration on Future Generations (Summit of the Future Outcome Documents). (2024). https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf

16   UNGA A/Res/70/1 Transforming Our World: The 2030 Agenda for Sustainable Development (2015). https://docs.un.org/en/A/RES/70/1

17   Rausch, C., & Luu, T. (2017). Inclusive Peace Processes Are Key to Ending Violent Conflict. US Institute of Peace. https://www.jstor.org/stable/resrep20175

18   O'Reilly, M., Ó Súilleabháin, A., and Paffenholz, T. (2015). "Reimagining Peacemaking: Women's Roles in Peace Processes," New York: International Peace Institute. https://www.ipinst.org/wp-content/uploads/2015/06/IPI-E-pub-Reimagining-Peacemaking-rev.pdf

19   Schädel, A., & Dudouet, V. (2020). Incremental inclusivity: A recipe for effective peace processes? Berghof Foundation. https://d1wqtxts1xzle7.cloudfront.net/65193162/BF_Incremental_inclusivity_final-libre.pdf?1608076243=&response-content-disposition=inline%3B+filename%3DIncremental_inclusivity_A_recipe_for_eff.pdf&Expires=1755381602&Signature=bSD0HzzPZCvdc1q-knWom76lVfiPcNbUs5xX1-Vkl-WoCzkrNGkqUHvkk9h9YRh4Yqfsvr9NaPLJH2N16FM92rizdKcQLLUDnFt4vI3AEkpBX7tJRoc95gfCN2f-3HBsx02pSOJK-FQk0Zp1qRsjcClNo0-ciVq9ZyjsHtKokdjLEoeUF2JcSg5DjtxUHuK1RecCBFVX5FMlvSgq-cLORTPwB4Lzn66RiNlvNiRxSXktlPQYqFjJVkq9HV-7LKfdWbCe-f6xzqcGcu87sZL1D~KoNuAmo7SxM793N-2ROky-YjqMOLdayDTlibZPpJ562ktKHlzU57wPBw2g7JG4Z5k51A__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

20   Harvard Kennedy School Center for Public Leadership – Negotiation and Conflict Resolution Collaboratory. (2024) Why It Worked: A research-based model for conflict transformation. See also: "Why It Worked: Redefining 'Success' in Conflict Resolution and Peace Negotiations" (2025) Giannone, M., Palmiano Federer, J., Kogutt, E., and Mandell, B. (eds.) Negotiation Journal Vol.41 Special Issue; available at: https://direct.mit.edu/ngtn/pages/ngtnv41si

21   Ghais, S. (2022). Inclusivity in Peace Processes: Civil Society and Armed Groups. In R. Mac Ginty & A. Wanis-St. John (Eds.), Contemporary Peacemaking: Peace Processes, Peacebuilding and Conflict (pp. 235–259). Springer International Publishing. https://doi.org/10.1007/978-3-030-82962-9_13

22   Palmiano Federer, J. (2024). The Trouble With Inclusivity: How Promoting Inclusive Peace Led to an Exclusive Outcome. In NGOs Mediating Peace: Promoting Inclusion in Myanmar's Nationwide Ceasefire Negotiations (pp. 153–177). Springer International Publishing. https://doi.org/10.1007/978-3-031-42174-7_7

23   Borger, J. (2015, November 10). Bosnia's bitter, flawed peace deal, 20 years on. The Guardian. https://www.theguardian.com/global/2015/nov/10/bosnia-bitter-flawed-peace-deal-dayton-agreement-20-years-on

24   Cuhadar, E. (2020). Understanding Resistance to Inclusive Peace Processes (No.159). United States Institute of Peace. https://www.usip.org/sites/default/files/2020-03/pw_159-understanding_resistance_to_inclusive_peace_processes-pw.pdf

25   UNGA Res A/78/L.49 | Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development, A/78/L.49 (2024). https://documents.un.org/doc/undoc/ltd/n24/065/92/pdf/n2406592.pdf

26   UN Security Council. (2023, July 18). UN Security Council S/PV.9381 9381st meeting, 18 July 2023, New York. Agenda: Maintenance of Peace and Security – artificial intelligence: Opportunities and risks for international peace and security. https://digitallibrary.un.org/record/4016424?v=pdf

27   UNGA. (2024, October 16). A/C.1/79/L.43 Artificial intelligence in the military domain and its implications for international peace and security | Seventy-ninth session. https://documents.un.org/doc/undoc/ltd/n24/299/16/pdf/n2429916.pdf

28   UN Department of Political and Peacebuilding Affairs (DPPA). (2024). DPPA Practice Note: Addressing the Malicious Use of Information and Communications Technologies in Mediation Processes. https://peacemaker.un.org/sites/default/files/document/files/2024/12/dppa-practice-note-addressing-malicious-use-icts-mediation-processes_1.pdf

29   UN Women. (2022). Action Brief – From Warfare to Peace building: Employing Artificial Intelligence for Women, Peace and Security. UN Women – Asia-Pacific. https://asiapacific.unwomen.org/en/digital-library/publications/2022/04/from-warfare-to-peacebuilding

30   Recommendation on the ethics of artificial intelligence. (2022). UNESCO. https://digitallibrary.un.org/record/4062376

31   Special Rapporteur on freedom of opinion and expression. (2018). A/73/348 Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment. https://www.ohchr.org/en/calls-for-input/report-artificial-intelligence-technologies-and-implications-freedom-expression-and#:~:text=In%20this%20report%2C%20the%20Special%20Rapporteur%20examines%20the,technologies%20comprising%20AI%20by%20states%20and%20private%20actors.

32   Boulanin, V., Brockmann, K., & Richards, L. (2020). Responsible Artificial Intelligence Research and Innovation for International Peace and Security. SIPRI. https://www.sipri.org/publications/2020/policy-reports/responsible-artificial-intelligence-research-and-innovation-international-peace-and-security

33   UNODA & SIPRI. (n.d.). Factsheet 4 | Responsible AI for peace and security. https://education.unoda.org/docs/ai4.pdf

34   Omowon, A. (2024). 'Technology and Conflict Resolution: Bridging Online and Offline Solutions.' International Journal of Research Publication and Reviews 5, no. 11 (January 2024): 7625–41. https://doi.org/10.55248/gengpi.5.1124.3430.

35   Home. (n.d.). VIEWS. https://viewsforecasting.org/

36   'VIEWS Data Dashboard'. Accessed 16 June 2025. https://data.viewsforecasting.org/.

37   'Martin Wählisch - Global Peace Lab'. Accessed 16 June 2025. https://globalpeacelab.net/

38   Study Group AI+X+G on "AI and the Future of Conflict Resolution: How Can Artificial Intelligence Improve Peace Negotiations?" Panel Discussion, Belfer Center, 16 April 2025.

39   Arana-Catania, M., Van Lier, F. A., and Procter, R. (July 23, 2022) 'Supporting Peace Negotiations in the Yemen War through Machine Learning'. arXiv. https://doi.org/10.48550/arXiv.2207.11528

40   Ibid.

41   Harvard Law School Program on Negotiation. (2025, March 9). PON AI Summit: AI in the Field [Video recording]. https://www.pon.harvard.edu/teaching-materials-publications/program-on-negotiation-ai-summit/

42   'Strategic Headwinds: Understanding the Forces Shaping Ukraine's Path to Peace | Projects | CSIS'. Accessed 11 May 2025. https://www.csis.org/programs/futures-lab/projects/strategic-headwinds-understanding-forces-shaping-ukraines-path-peace.

43   AI models could help negotiators secure peace deals. (2025, April 16). The Economist. https://www.economist.com/science-and-technology/2025/04/16/ai-models-could-help-negotiators-secure-peace-deals

44   Williams, R. (October 17,. 2024). AI could help people find common ground during deliberations. MIT Technology Review. https://www.technologyreview.com/2024/10/17/1105810/ai-could-help-people-find-common-ground-during-deliberations/

45   Davis, N. (2024) 'AI Mediation Tool May Help Reduce Culture War Rifts, Say Researchers'. The Guardian, 17 October 2024, sec. Technology. https://www.theguardian.com/technology/2024/oct/17/ai-mediation-tool-may-help-reduce-culture-war-rifts-say-researchers

46   "United Nations Department of Political and Peacebuilding Affairs, Innovation in Peacebuilding: Technology, Data and Digital Tools in Conflict Prevention and Peacebuilding" (New York: United Nations, 2022), https://dppa.un.org/sites/default/files/innovation.pdf; see also Using the Power of AI to Boost Youth Political Participation in Guatemala | Department of Political and Peacebuilding Affairs. (2024, December 3). https://dppa.un.org/en/using-power-of-ai-to-boost-youth-political-participation-guatemala, and United Nations Department of Political and Peacebuilding Affairs & and Centre for Humanitarian Dialogue. (2019, March). DIGITAL TECHNOLOGIES AND MEDIATION IN ARMED CONFLICT. https://www.hdcentre.org/wp-content/uploads/2019/07/Digital-Technologies-and-Mediation-in-Armed-Conflict.pdf

47   "United Nations System Chief Executives Board for Coordination, Report on the Operational Use of AI in the UN System (November 2024), 24" https://unsceb.org/sites/default/files/2024-11/Report%20on%20the%20Operational%20Use%20of%20AI%20in%20the%20UN%20System_1.pdf

48   'Global Conference on AI, Security and Ethics 2025 → UNIDIR', 7 January 2025. https://unidir.org/event/global-conference-on-ai-security-and-ethics-2025/. See Report: Anand, A., & He, W. (2023). The 2022 Innovations Dialogue: AI Disruption, Peace and Security | Conference Report. UNIDIR. https://unidir.org/wp-content/uploads/2023/09/UNIDIR_ID22_AI_Disruption_Peace_Security_Conference_Report.pdf

49   Albrecht, E., Fournier-Tombs, E., & Brubaker, R. (2024). Disinformation and Peacebuilding in Sub-Saharan Africa: Security Implications of AI-Altered Information Environments. United Nations University. https://unu.edu/sites/default/files/2024-11/disinformation%20peacebuilding%20subsaharan%20africa.pdf

50   UNGA Res A/78/L.49 | Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development, A/78/L.49 (2024). https://documents.un.org/doc/undoc/ltd/n24/065/92/pdf/n2406592.pdf

51   'Inter-Agency Working Group on Artificial Intelligence | United Nations - CEB'. Accessed 16 June

2025. https://unsceb.org/inter-agency-working-group-artificial-intelligence

52   Artificial Intelligence and its impact on peace, security and governance – Amani Africa. (2025, March 19). https://amaniafrica-et.org/artificial-intelligence-and-its-impact-on-peace-security-and-governance/

53   OECD. 'Global Partnership on Artificial Intelligence'. Accessed 16 June 2025. https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html

54   The Association of Southeast Asian Nations (ASEAN). (2024). ASEAN Guide on AI Governance and Ethics. https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf

55   The Association of Southeast Asian Nations (ASEAN). (2025). ASEAN Digital Masterplan 2025. https://asean.org/wp-content/uploads/2021/09/ASEAN-Digital-Masterplan-EDITED.pdf

56   The Association of Southeast Asian Nations (ASEAN). (2024, October 11). ASEAN-U.S. LEADERS' STATEMENT ON PROMOTING SAFE, SECURE, AND TRUSTWORTHY ARTIFICIAL INTELLIGENCE. https://asean.org/wp-content/uploads/2024/10/11.-2024-ASEAN-US_Leaders_Statement_on_AI_FINAL.pdf

57   GESDA Global. 'Home'. Accessed 16 June 2025. https://gesda.global/

58   Ministry of Foreign Affairs, Republic of Korea. (2025, February 26). Ministry of Foreign Affairs Launches AI Diplomacy Division. https://www.korea.net/Government/Briefing-Room/Press-Releases/view?articleId=322749&type=N&insttCode=A260104

59   Allen, G. C., and Adamson, G. (October 30, 2024) 'The AI Safety Institute International Network: Next Steps and Recommendations'. https://www.csis.org/analysis/ai-safety-institute-international-network-next-steps-and-recommendations

60   Campus France. 'The French Government Creates a National Institute to Assess and Secure AIs', 14 June 2025. https://www.campusfrance.org/en/actu/creation-d-un-institut-national-pour-l-evaluation-et-la-securite-de-l-ia

61   Building a responsible humanitarian approach: the ICRC's policy on artificial intelligence | International Committee of the Red Cross', 18 November 2024. https://www.icrc.org/en/publication/building-responsible-humanitarian-approach-icrcs-policy-artificial-intelligence

62   Thompson, S., & Piirtola, A. (2024). ARTIFICIAL INTELLIGENCE AND PEACEMAKING: The case of digital dialogues in Sudan. CMI Insight | CMI-Martti Ahtisaari Peace Foundation. https://cmi.fi/wp-content/uploads/2024/02/CMI_INSIGHT_2024_sudan.pdf

63   AI for Peace. 'Ethics'. Accessed 16 June 2025. https://www.aiforpeace.org/ethics

64   NL AI Coalition. (2022). AI Applications for Security, Peace and Justice: Effective, secure and human centric | NL AI Coalition Working Group Security, Peace and Justice. https://nlaic.com/wp-content/uploads/2022/02/AI-Applications-for-Security-Peace-and-Justice.pdf

65   https://www.csis.org/programs/futures-lab Futures Lab | Defense and Security | CSIS. https://www.csis.org/programs/futures-lab; and Artificial Intelligence and International Stability Project. CNAS. https://www.cnas.org/artificial-intelligence-and-international-stability-project

66   RAND. Artificial Intelligence. https://www.rand.org/topics/artificial-intelligence.html

67   AI Across Sectors: Bridging Technology and Peacebuilding | NYU Steinhardt. Accessed 16 June 2025. https://steinhardt.nyu.edu/ihdsc/ai-across-sectors-bridging-technology-and-peacebuilding

68   Belfer Center's AI + X + G Seminar- AI and Conflict Resolution | The Belfer Center for Science and International Affairs, 16 April 2025. https://www.belfercenter.org/event/belfer-centers-ai-x-g-seminar-ai-and-conflict-resolution

69   Oakes, V. 'How a People-First Approach to AI and Technology Can Help Empower Humanitarian Action'. Microsoft Industry Blogs - United Kingdom, 19 November 2020. https://www.microsoft.com/en-gb/industry/blog/cross-industry/2020/11/19/ai-for-humanitarian-action/

70   Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., and De Beer, J. (2020). 'Artificial Intelligence (AI) Deployments in Africa: Benefits, Challenges and Policy Dimensions'. The African Journal of Information and Communication (AJIC), no. 26 (15 December 2020). https://doi.org/10.23962/10539/30361.

71   Hirblinger, A. T. (2022) 'When Mediators Need Machines (and Vice Versa): Towards a Research Agenda on Hybrid Peacemaking Intelligence'. International Negotiation 28, no. 1 (25 January 2022): 94–125. https://doi.org/10.1163/15718069-bja10050..

72   'ASRSG Williams Conducts Digital Dialogue with 1000 Libyans | Department of Political and Peacebuilding Affairs'. Accessed 16 June 2025. https://dppa.un.org/en/asrsg-williams-conducts-digital-dialogue-with-1000-libyans#:~:text=ASRSG%20Williams%20conducts%20digital%20dialogue,of%20Political%20and%20Peacebuilding%20Affairs.

73   Omawon, A. (2024).

74   Mäki, N. (2020) 'Between Peace and Technology – A Case Study on Opportunities and Responsible Design of Artificial Intelligence in Peace Technology', 1 January 2020. https://www.academia.edu/88332979/Between_Peace_and_Technology_A_Case_Study_on_Opportunities_and_Responsible_Design_of_Artificial_Intelligence_in_Peace_Technology.

75   'UN Hires an AI Company to Analyse the Israeli-Palestinian Conflict | Digital Watch Observatory', 3 November 2023. https://dig.watch/updates/un-hires-an-ai-company-to-analyse-the-israeli-palestinian-conflict.

76   Lee, S. (2025, May 25). Peacebuilding in the Digital Era. https://www.numberanalytics.com/blog/peacebuilding-in-the-digital-era

77   Ebongue, J. (2025) 'From Community Network to Community Data: Towards Combining Data Pool and Data Cooperative for Data Justice in Rural Areas'. arXiv, 7 March 2025. https://doi.org/10.48550/arXiv.2503.05950.

78   Thu, H. D. (2025). 'AI Bias and Fairness: The Definitive Guide to Ethical AI'. SmartDev (blog), 15 April 2025. https://smartdev.com/addressing-ai-bias-and-fairness-challenges-implications-and-strategies-for-ethical-ai/.

79   Mandokhail, A. W. K. (2024). 'The Transformative Role of Artificial Intelligence in Conflict Resolution and Peacekeeping'. NUST Journal of International Peace & Stability, 31 January 2024, 104–9. https://doi.org/10.37540/njips.v7i1.166.

80   Hansen, A. S. (2023). 'Peacekeeping in the Digital Age: Future Threats and Capability Requirements'. In The EU, Irish Defence Forces and Contemporary Security, edited by Jonathan Carroll, Matthew G. O'Neill, and Mark Williams, 229–49. Cham: Springer International Publishing, 2023. https://doi.org/10.1007/978-3-031-07812-5_12..

81   Trusilo, D., and Danks, D. (2024) Commercial AI, Conflict, and Moral Responsibility: A theoretical analysis and practical approach to the moral responsibilities associated with dual-use AI technology. arXiv, 30 January 2024. https://doi.org/10.48550/arXiv.2402.01762.

82   Pizzi, M., Romanoff, M., and Engelhardt, T. (2020). 'AI for Humanitarian Action: Human Rights and Ethics'. International Review of the Red Cross 102, no. 913 (April 2020): 145–80. https://doi.org/10.1017/S1816383121000011; see also Latonero, M. (2019). "Opinion | Stop Surveillance Humanitarianism." The New York Times, July 12, 2019, sec. Opinion. https://www.nytimes.com/2019/07/11/opinion/data-humanitarian-aid.html.

83   "Press Release | AI Action Summit: Ensuring the Development of Trusted, Safe and Secure AI to Benefit of All," February 12, 2025. https://www.sgdsn.gouv.fr/files/2025-02/20250211_Press%20release_AI%20Summit_Trust_ENG_1.pdf; also Renaissance numérique. "10 AI Governance Priorities: Survey of 40 Civil Society Organisations." https://www.renaissancenumerique.org/en/publications/10-ai-governance-priorities-survey-of-40-civil-society-organisations/.

84   Our History | Peace Appeal'. Accessed 16 June 2025. https://peaceappeal.org/history/

85   Study Group AI+X+G on "AI and the Future of Conflict Resolution: How Can Artificial Intelligence Improve Peace Negotiations?" Panel Discussion, Belfer Center, 16 April 2025.

86   Nowotny, H. (2024). AI and the illusion of control. Proceedings of the Paris Institute for Advanced Study, 2024, Vol.1, ff10.5281/zenodo.13588581ff. ffhal-04706414f https://hal.science/hal-04706414v1/file/ai-and-the-illusion-of-control.pdf

87   Czaja, J. (2024). 'The Application of Artificial Intelligence-Based Decision-Assisting Tools to the Mediation Process – an Analysis of Risk and Opportunities'. Forum Prawnicze., no. 4 (84) (2024). https://doi.org/10.32082/fp.4(84).2024.1296.

88   Ahmad, N. (2021). Refugees and Algorithmic Humanitarianism: Applying Artificial Intelligence to RSD Procedures and Immigration Decisions and Making Global Human Rights Obligations Relevant to AI Governance. International Journal on Minority and Group Rights 28 (2021): 367–435. https://doi.org/10.1163/15718115-bja10007.

89   Possati, L. (2024). Mediation and Anti-Mediation: Re-Evaluating the Role of AI Breakdowns and Anomalies in Postphenomenology. Journal of Human-Technology Relations. 2. 10.59490/jhtr.2024.2.7381. https://www.researchgate.net/publication/385672118_Mediation_and_Anti-Mediation_Re-Evaluating_the_Role_of_AI_Breakdowns_and_Anomalies_in_Postphenomenology

90   Kasapoğlu, C. and Kırdemir, B. "Wars of None: Artificial Intelligence and the Future of Conflict." Centre for Economics and Foreign Policy Studies, 2019. http://www.jstor.org/stable/resrep21050.

91   Engelke, P. (2020). "AI, Society, and Governance: An Introduction." Atlantic Council, 2020. http://www.jstor.org/stable/resrep29327.