IPIE
International Panel on the Information Environment

Artificial Intelligence and Peacebuilding

Opportunities and Challenges

Technical Paper 2025.3

Artificial Intelligence and Peacebuilding

Opportunities and Challenges

Technical Paper 2025.3

How to cite:
International Panel on the Information Environment [C. Zelizer, F. Ogenga, L. Schirch, E. Tauchnitz, P. N. Howard, S. Valenzuela (eds.)], "Artificial Intelligence and Peacebuilding: Opportunities and Challenges," Zurich, Switzerland: IPIE, 2025. Technical Paper, TP2025.3, doi: 10.61452/RNGW7145.

SYNOPSIS

What role can artificial intelligence (AI) play in building peace? Can AI technology help understand and monitor conflict emergence, assist diverse stakeholders' responses to violence, and help in reconciliation and post-conflict reconstruction? While these questions have been part of peacebuilding discussions for years, today's AI capabilities offer unprecedented possibilities.

Though some applications—particularly in conflict mapping and analysis—are gradually entering mainstream practice, most peacebuilding AI initiatives remain in pilot phases. Current experimentation spans conflict prediction systems, AI-facilitated public dialogues, enhanced crisis response tools, and solutions that improve organizational efficiency—presenting both opportunities and challenges for peacebuilding.

While AI demonstrates valuable capabilities for peacebuilding, it remains fundamentally a dual-use technology. The same systems that can better predict conflict hotspots, help facilitate dialogue between opposing groups, and process humanitarian data can efficiently spread misinformation, amplify division, and deepen societal polarization. The technology's substantial energy requirements and surveillance applications present additional concerns.

AI tools in conflict zones provide a means for tracking violence, delivering intelligence to security actors including peacekeepers, and potentially protecting civilians through more effective peacekeeping operations. However, these capabilities can create more harm: fully autonomous weapons operating without human oversight; surveillance technologies enabling government oppression; and "smart" targeting systems potentially increasing civilian casualties.

This Technical Paper outlines how AI can aid peacebuilding amid current global challenges including violent conflict, inequality, climate change, resource scarcity, and forced migration. It reviews the current state of AI and peacebuilding research and practice, and related fields including humanitarian action and international development, and offers recommendations to better harness AI's potential for peace.

Core recommendations for ensuring inclusive and ethical AI development include: (1) prioritizing fundamental human rights; (2) increasing funding for applied practices that document effective and ineffective strategies; (3) supporting ecosystem organizations to foster cross-sectoral collaboration; (4) encouraging the development of smaller language models; and (5) developing strong policy and ethical guidance firmly anchored in human rights to promote ethical AI use and mitigate potential risks or harms.

CONTENTS

SECTION 1. INTRODUCTION

The current widespread digital transformation, often termed the Fourth Industrial Revolution, has been reshaping societies over the past decade, bringing unprecedented challenges and opportunities. New forms of big data, robotics, the Internet of Things, and growing computational power have driven many such transformations. Artificial intelligence (AI) is currently being used to address a wide range of development challenges, and both quantum computing and nanotechnology may provide additional opportunities in the years ahead [1], [2]. At its core lies data—the capacity to collect, analyze, and leverage information built on data to influence decision-making across sectors. While data may seem neutral, the methods of its collection, the entities controlling it, and the purposes driving its use profoundly shape its impact. Data exists within complex information environments that influence how it emerges, disseminates, and creates value—determining who benefits and who faces exclusion in an increasingly data-driven society [3], [4].

Data is being widely used in the emerging information economy with dual effects: it can contribute to positive social development while also fueling polarization, undermining democracy, and deepening surveillance. How data is used in a society depends to a significant degree on the wider information environment. In more open environments—where there is free media, public access to information, strong data governance, effective technology regulation, and high digital literacy—AI tools and data may help foster more resilient and peaceful societies. While the weaponization of AI tools still occurs in democracies, as evidenced by numerous global examples, these societies may be more resistant to such challenges. Conversely, fragile or closed environments are often marked by censorship, disinformation, weak institutions, and limited civic spaces. In those environments, the same technologies are even more suitable for weaponization, possibly reinforcing control, limiting civic engagement, and fueling conflicts [5].

What role can AI play in building peace? Can AI technology help understand and monitor conflict emergence, assist diverse stakeholders' responses to violence, and help in reconciliation and post-conflict reconstruction? AI and peacebuilding cannot be treated as isolated fields; their positive or

negative impacts are inherently connected to the larger political, economic, and information environments in which they operate [6], [7]. This Technical Paper looks at AI and peacebuilding within this larger context with a focus on the information environment.

SECTION 2. METHODOLOGY

The Scientific Panel on AI and Peacebuilding, which is part of the International Panel on the Information Environment, produced this Technical Paper. To best understand the state of this nascent field, a wide variety of literature was consulted and a broad search for the latest cases was conducted. Peer-reviewed articles, policy and research reports by think tanks, civil society, funders, academics, and the private sector, as well as news articles and selected podcasts and videos were reviewed.

Over 600 documents were reviewed, categorized, and explored for common themes, case studies, ethical and policy recommendations, and future trends using search parameters on a closed corpus. Efforts were made to include a wide range of literature from diverse sources, perspectives, and geographic locations. The documents were uploaded into Logically AI, an AI-powered research platform, and reviewed and classified by thematic focus. A customized closed dataset was then created and analyzed using extensive Large Language Model (LLM) prompting. Particular keywords and relevant databases were searched intensively (see Annex).

This Technical Paper is organized into seven main sections:

1. **Development of Artificial Intelligence**—Key aspects and the evolution of artificial intelligence;
2. **Building Safer AI: Ethics, Power, and Rights**—Potential dangers and misuse of AI. Importance of integrating human rights into the design of AI to ensure AI serves human needs safely;
3. **Overview of AI and Peacebuilding**—Overview of the integration of the fields;
4. **Applied Use of AI Tools**—Evidence of AI implementation across conflict prevention, humanitarian response, civic engagement, operations, and training;
5. **Building a Healthy Information Environment**—Characteristics of a healthy information environment and AI's role in supporting this;

6. **Future Areas of Opportunity**—Emerging opportunities and potential challenges at the AI-peacebuilding intersection; and
7. **Key Recommendations and Conclusion**—Practical guidance for responsibly leveraging AI for peace while minimizing risks.

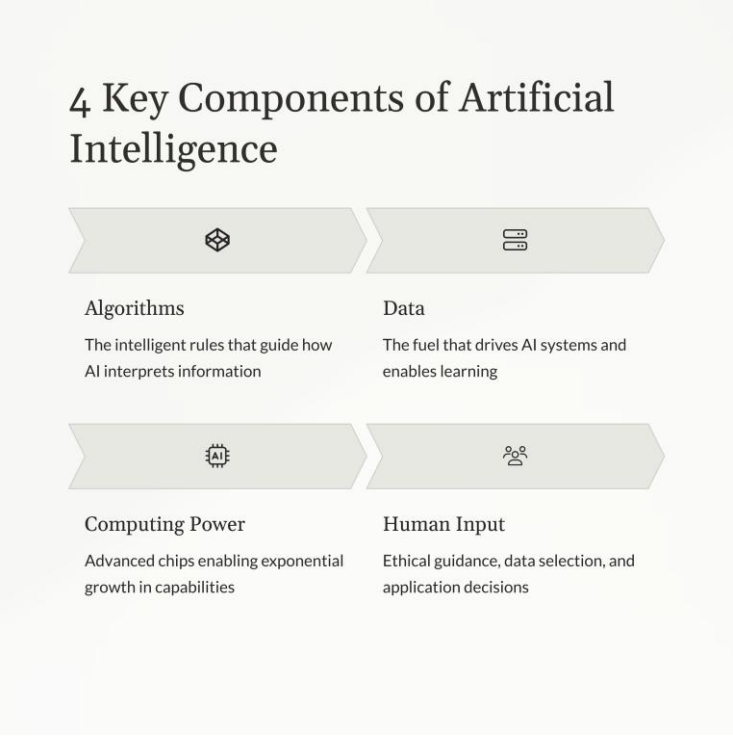
SECTION 3. THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

Since the concept of modern AI emerged in the 1950s, AI has been envisioned as a futuristic technology that would someday be integrated into everyday life and society [8]. That envisioned future may now be arriving, with AI embedded in everything from data analysis and digital assistance to decision support in finance, education, and governance.

While AI lacks a universally agreed upon definition, writer Madhumita Murgia provides a useful framework for use in this Technical Paper describing AI as "a complex statistical software applied to finding patterns in large sets of real-world data" [p.4, 9]. At its core, AI uses algorithms to interpret complex data at scales and speeds beyond human capacity [10].

Figure 1 offers a visualization of four key components that enable AI functionality, based on several valuable assessments from the literature. The figure draws on work by Ramanathan & Fruchterman, Abbott & Elliot, and Mansell et al. [11], [12], [13].

Figure 1. Four Key Components of AI



Source: Drawing upon Ramanathan & Fruchterman, Abbott & Elliot and Mansell et al.

Over the past decade, powerful innovations have led to the emergence of generative AI. Early AI systems operated on simple rule-based logic with predetermined outputs and confined parameters. These basic systems, including decision trees and scripted chatbots, could not learn beyond their initial programming [14], [15]. The paradigm shifted with the growth of machine learning, which evolved from simple statistical models to sophisticated systems capable of learning patterns from the data itself [16]. As Cuéllar et al. note, "it was much harder to program a computer to be clever than it was to program a computer to learn to be clever" [p.5, 17]. The true breakthrough came with deep learning via neural networks which have radically scaled pattern recognition, classification and predictive modelling [16], [18].

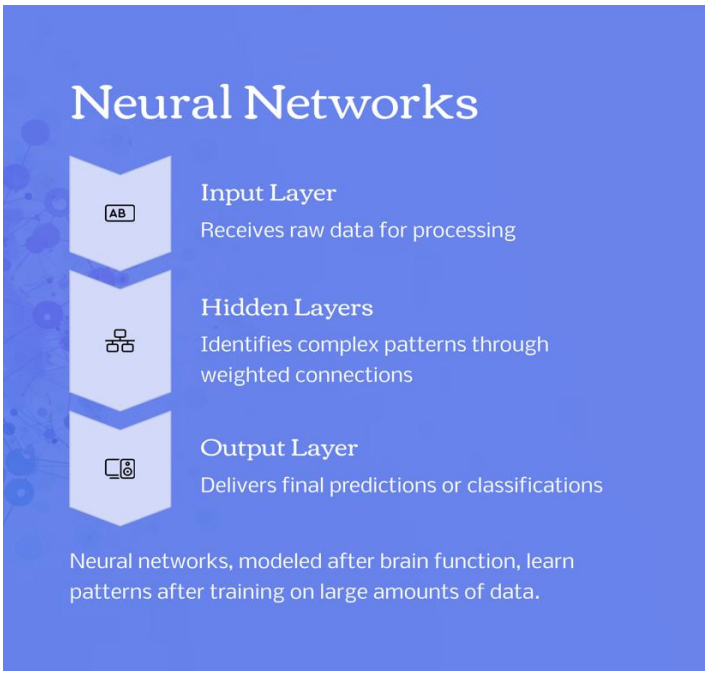
Foundation models represent a significant advancement in AI technology. Reichstein et al. define a foundation model as "a large deep neural network trained in a self-supervised manner on a large body of unlabeled data, to subsequently enable its application to many different (downstream) tasks" [p.6, 19]. Large language models (LLMs), among the first foundation models, analyze patterns across extensive datasets to produce human-like text beyond their original training parameters. LLMs belong to the broader category of generative AI— technologies that produce text, images, video, and other media using similar algorithmic processes and training methodologies [13], [20].

Neural networks, which are computational systems modeled after brain function, drive AI's exponential growth. These networks range from basic architectures to complex, deep neural networks with multiple layers. Processing typically follows three stages: data enters the input layer for formatting, passes through "hidden layers" where pattern recognition occurs, and reaches the output layer to produce results based on training [18].

This process is outlined below in figure 2, Overview of Neural Networks. The figure draws from work from Linardos and Schirch [18]. [1]

[1] Lisa Schirch, "Public Discourse and LLMs," PowerPoint presentation, personal communication with Craig Zelizer, March 15, 2025.

Figure 2. Overview of Neural Networks



Source: Drawing upon Linardos and Schirch.

Prompt engineering is currently the primary method for human–AI interaction. Prompting techniques range in complexity from simple questions to "few-shot" prompting that provides examples of the type of output the user seeks. In contrast, "chain-of-thought" prompting guides the AI system through a reasoning process by providing context and explanation to increase the relevance and quality of the output [21], [22]. These approaches differ mainly in how much context and guidance is provided to the model, with more detailed prompts typically producing better outputs. Learning how to do effective prompting is a skill that can be advanced through better understanding how LLMs function and through extensive practice.

Current AI models employ both supervised and unsupervised learning. Supervised learning uses labeled data to help algorithms identify patterns [18], [23], [24]. This work is often conducted by data labelers or annotators working under supervision. It is an iterative process involving fine-tuning data, algorithms, and testing, and having data that is as reliable and unbiased as possible is critical. Unsupervised learning relies on unlabeled datasets analyzed by algorithms to discover patterns and

similarities [23], [25]. Machine learning processes powered by algorithms analyze the data to discover patterns, grouping, and similarities [25], [26]. It is also possible to combine both approaches for semi-supervised learning, training on smaller data sets and then looking for patterns in larger amounts of unlabeled data [25], [27].

An overview of the two approaches is provided in Table 1. The table draws on work by Croicu, Hawke et al., and Mueller and Rauh, regarding applied examples of peacebuilding data [16], [28], [29].

Table 1. Training Supervised and Unsupervised Models

Feature	Supervised learning	Unsupervised learning
Training data	Labeled data (e.g., news articles labeled as indicating conflict or peace)	Unlabeled data (e.g., social media posts, news articles, radio transcripts)
Learning goal	Predict outcomes (e.g., conflict likelihood) based on labeled examples	Discover patterns, structures, and relationships in data without predefined labels
Peacebuilding examples	Conflict prediction using labeled datasets of conflict events and indicators; text classification to identify hate speech or polarization	Topic modeling identifies themes in discussions about conflicts on social media, grouping similar disputes to find common issues

Source: Drawing upon Croicu, Hawke et al., and Mueller and Rauh.

Despite advances, even the best generative AI systems still require significant human involvement in data selection, labeling, training, variable weighting, testing, and safety implementation [28], [30], [31]. Though AI output may seem magical and human-like, these systems are sophisticated prediction machines, not truly intelligent entities. Most data labeling for LLMs is performed by workers in low- and middle-income countries who are often working under precarious conditions with inadequate compensation [32], [33]. The technical expertise required for building and deploying models demands highly skilled engineers and machine learning specialists. Despite improvements in training and accuracy, AI outputs still contain inaccuracies and hallucinations that go beyond inaccurate to fabricated content, though at decreasing rates [34], [35].

The potential emergence of Artificial General Intelligence (AGI) marks a critical threshold beyond today's AI tools.[2] Experts disagree on how to define this milestone, though many look for human-equivalent creativity and cross-domain problem-solving capabilities [36], [37], [38]. Predictions for AGI development span from years, to decades, to never. The achievement of AGI will likely significantly impact peacebuilding with some important limitations. Though AI may model emotional intelligence, it is unlikely to be able to replicate the processes needed to advance human trust-building essential to peacemaking. Conflicts emerge from complex identity dynamics and systemic inequalities, with resolution requiring relationship transformation rooted in human agency. AGI will likely be an invaluable analytical and process tool—mapping conflicts, outlining policy and programmatic options, and forecasting potential outcomes—yet authentic trust-building involves cultural nuances challenging for algorithms to navigate. While AGI may enhance peacebuilders' work through pattern recognition and process optimization, the human engagement that fosters relational transformation for positive peace will likely be hard to substitute with technology [39].

[2] A popular framework for the emergence of more advanced AI is first artificial narrow intelligence (ANI), followed by artificial general intelligence (AGI)—and some futurists then envision a future with artificial superintelligence [36], [37].

SECTION 4. BUILDING SAFER AI: ETHICS, POWER, AND RIGHTS

AI tools offer numerous positive applications in peacebuilding, education, health, economics, and governance. However, these same technologies are simultaneously being deployed for harmful purposes: increased surveillance, repression, polarization, political manipulation, and information distortion [40], [41]. For AI to genuinely benefit peacebuilding efforts, human rights principles such as respect for privacy, protection of civil and political rights and liberties, personal freedoms, and accountability and safety must be integrated throughout the entire lifecycle of AI systems, from design and development to assessment [42], [43]. Even well-intentioned applications can undermine human rights and harm vulnerable communities in conflict zones if these considerations are overlooked [41], [44]. The United Nations issued a set of principles around the ethical use of AI that integrates the core principles outlined above as well as adding others, such as the need for human oversight and factoring in sustainability [43].

With an ever-growing prevalence of technology across societies, centering ethics has become imperative. Many practitioners and scholars increasingly advocate embedding ethical frameworks across all AI development and deployment phases [45], [46], [47]. Despite the proliferation of ethical guidelines for AI use, including UNESCO's recommendations on the use of AI, the G20 Principles for Responsible Stewardship of Trustworthy AI, the EU AI Act and dozens of others, ethics-washing has become a common phenomenon that powerful technology corporations can employ to circumvent the legal obligations imposed by existing human rights law [5], [39], [48]. These approaches, whether labeled ethical, responsible or compassionate AI, share core principles: do no harm, respect human rights, integrate ethical governance, employ informed consent, and ensure meaningful human oversight.

As Tauchnitz has pointed out, applying an "ethics of human rights" approach to AI and peacebuilding that is firmly grounded in international human rights norms can serve as a point of reference in the increasingly complex ethical and legal landscape surrounding peace and conflict [45]. If human rights were strictly applied both online and offline, arguably there would be little need for additional guidelines for AI [49].

Centering ethics and human rights in peacebuilding is crucial for several reasons. First, populations affected by conflict typically exist in vulnerable economic, cultural and political circumstances that may exacerbate their marginalization [50], [51]. In heightened conflict situations, AI tools designed to address disinformation and hate speech can paradoxically help facilitate direct violence against identity groups, resulting in substantial harm to both people and physical infrastructure [50], [51]. Additionally, data privacy becomes vital in these contexts, as data breaches or misuse can lead to direct targeting of vulnerable populations—whether through inadvertent exposure or deliberate weaponization of these technologies against specific communities [52].

Power and money play a key role in shaping AI ethics and safety. Because AI development is largely funded by venture capital, governments, and military interests, financial incentives often overshadow ethical considerations. Despite rhetoric about AI's potential to benefit humanity through sectors like education and healthcare, investors primarily seek financial returns. Companies that publicly champion ethical values frequently prioritize profits for a small group of stakeholders [17], [53]. As an AI arms race accelerates globally, there is a risk that data security, ethics, and the public good are sacrificed for competitive advantage and the benefits of scale [46], [54], [55]. This echoes the Cold War era, when nations competed for military dominance—only now, corporations have joined governments in the sprint for technological supremacy [54], [56], [57].

Civil society organizations, multilateral bodies, public interest groups, and other stakeholders actively shape how AI ethics are conceptualized and implemented through various channels. These include public–private partnership groups, research initiatives, international forums, and rights advocacy campaigns. Notable examples include the UN Summit for the Future (September 2024), which developed guiding principles for our digital future. Similarly, the EU Safety Act, which established comprehensive regulatory safeguards, and the African Union's movement toward a continent-wide approach to AI regulation and governance illustrate how coalitions around AI ethics have formed [58]. Organizations such as the Distributed AI Research Institute have also contributed valuable rights-based perspectives to these discussions. However, it remains crucial to acknowledge

that, outside the EU's robust regulatory framework, significant power imbalances exist between big technology companies and these civil society initiatives [59], [60].

Algorithms and autonomous weapons are increasingly used in targeting and warfare in combat, creating new legal challenges and accountability gaps in the case of human or technical errors [61]. While humans are the ones making final targeting decisions, this may change in the near future. In Ukraine and Gaza, AI has played an important role in targeting decisions, often with horrific outcomes for civilians and infrastructure. The Israeli military has deployed several AI models that have been critical in targeting recommendations. One system, Gospel, can generate hundreds of targeting recommendations within a week—a fraction of the time and human capital previously needed [62], [63], [64]. The Israel Defense Forces use a parallel AI system, Lavender, to mark individuals linked to Hamas or other militant groups and to make recommendations for targeted killing [63], [65], [66]. Ukraine and Russia both use drones widely in attacks, and semiautonomous drones are becoming more prominent [67], [68].

In addition to the concerns raised around environmental impact, disinformation, and military and safety risks, another key challenge is that current LLMs hallucinate and frequently produce inaccurate outputs. Research from Vectra AI, an AI cybersecurity firm, shows even the best current generation LLMs hallucinate between 0.8% and 5% of the time [35], [69]. These error rates, measured in ideal conditions, almost certainly worsen in places with limited information or complex situations—exactly the environments where humanitarian work and conflict response occur. While a few mistakes might not matter for everyday uses, they become dangerous when lives and critical operations depend on accurate information.

Effective safety and governance frameworks are essential for ethical AI use. Although governments and key stakeholders are developing and implementing AI policies, most remain voluntary rather than legally binding [58]. The 2024 UN report, *Governing AI for Humanity* emphasizes that "human rights must be at the centre of AI governance, ensuring rights-based accountability across jurisdictions" [p.39, 70]. This urgency is reinforced by a recent survey conducted by the International

Panel on the Information Environment, which found that 63% of over 400 AI experts predict a deterioration of the information environment in the coming year [7].

SECTION 5. OVERVIEW OF AI AND PEACEBUILDING

A key challenge at the intersection of AI and peacebuilding involves conceptual clarity, as both fields suffer from significant terminological ambiguity. There is a distinction between peacebuilding as a set of practical tools, policies, and interventions, in contrast with peace and conflict studies as an established academic discipline [71]. For this Technical Paper, peacebuilding is defined as supporting people in or at risk of conflict to prevent or end direct violence by addressing structural, policy, and relational differences while strengthening institutions and spaces that foster peaceful relations [71], [72]. This approach would necessarily require evidence generation through research for programmatic decision-making, bridging practical applications with the academic realm. Despite varying terminology across these domains, significant commonalities emerge in AI applications spanning civic participation, public deliberation, democratic dialogue, social cohesion technologies, and violence prevention efforts [73], [74].

This report adopts the UN's triple nexus framework, which recognizes the interconnected nature of peace, humanitarian, and development efforts [49], [71], [75]. Since the 1990s, this integrated approach has gained traction as peacebuilding principles have been increasingly incorporated into development and humanitarian programming—reflecting the reality that much of this work occurs in conflict-affected contexts requiring coordinated responses [76]. A common operational component within this framework is the deployment of peacekeepers, police, and civilian experts. These personnel can be deployed during highly complex negotiations between or within states, or after agreements when implementing ceasefires, monitoring conditions, and providing security for civilian populations [77]. This framework acknowledges peacebuilding both as a standalone practice and as an essential component within broader humanitarian and development initiatives, offering practical tools for understanding conflict dynamics, analyzing key stakeholders, strengthening societal resilience, and facilitating post-conflict reconciliation.

Central to effective peacebuilding are principles such as "do no harm" and conflict sensitivity, developed by Mary Anderson and colleagues at CDA Collaborative Learning Projects. This approach

recognizes that even well-intentioned interventions can exacerbate tensions in resource-scarce environments by influencing "connectors" that help build peace, and "dividers" that can increase polarization [78]. While external actors can support peace processes, sustainable impact ultimately depends on centering local knowledge and actors. It is also critical to strengthen local institutions and capacity—an approach now essential across humanitarian aid, development, and other sectors working in conflict-affected regions.

The peacebuilding field encompasses diverse stakeholders working at multiple levels: civil society organizations, local and international NGOs, government agencies, faith-based groups, educational institutions, think tanks, multilateral agencies (including the UN and its specialized bodies), bilateral donors, and private sector entities. Thousands of organizations worldwide apply their distinct expertise to conflict analysis, violence prevention, crisis response, and community resilience-building. Supporting these efforts, specialized networking organizations have emerged to facilitate knowledge-sharing, policy advocacy, and strategic collaboration. Key examples include the African Peacebuilding Network, West Africa Network for Peacebuilding, European Peacebuilding Liaison Office, and global platforms such as the Global Partnership for the Prevention of Armed Conflict and the Alliance for Peacebuilding.

Multilateral organizations increasingly use digital tools to strengthen peacebuilding capacity, learning and collaboration. A notable example is the African Union, which is leveraging its regional expertise to launch a digital Massive Open Online Course (MOOC) in 2025, designed to enhance peacebuilding capacity among civil society organizations across Africa.[3]

Peacebuilding funding remains modest within the broader international development landscape. According to the Organisation for Economic Co-operation and Development's Development

[3] Interview with Dr. Gotwam Raj Chintaram, Program Officer, CSO Engagement, Africa Union ECOSOCC, conducted by Dr. Fredrick Ogenga, April 26, 2025. For more on the MOOC see https://ecosocc.au.int/en/news/press-releases/2024-12-18/new-leadership-apsa-and-peace-and-security-cso-database

Assistance Committee (OECD-DAC), the amount of official development assistance (ODA) given by its members to high and extreme fragility contexts totaled just 5.3 billion USD in 2023, with only 1.7 billion USD specifically allocated to conflict prevention activities [79]. UN Peacekeeping operations, an essential part of many peace processes, had a 2024 budget of 5.6 billion USD, while global military expenditure reached 2.443 trillion USD [80], [81]. The funding outlook appears increasingly bleak: USAID has been devastated by recent budget cuts and total elimination of the agency, alongside reductions from at least nine other OECD bilateral donors. Projections indicate international development assistance by OECD-DAC members may decline by approximately 25% over the next two years compared to 2023 levels [82]. These cuts are a risk to operational innovation and the application of new technologies to development and peacebuilding broadly.

For over two decades before AI's emergence in peacebuilding, practitioners leveraged diverse technologies—drones, satellites, big data, crisis mapping, social media, and online dialogue platforms. Importantly, technology integration is not novel to this field. Today's organizations pioneering AI applications build upon established pre-generative AI innovations. This technological evolution in peacebuilding will be examined in subsequent sections.

Mapping to Date

The literature on AI and peacebuilding remains largely conceptual, focusing on potential applications rather than on widespread practical implementation with a proven impact. Although documented AI deployments in peacebuilding are increasing, strong evidence of measurable outcomes is limited.

AI tools are increasingly being integrated into peacebuilding work by organizations worldwide. Leading this movement is the UN's International Telecommunication Union (ITU) through its AI for Good platform, established in 2017 in collaboration with various UN agencies and stakeholders [83]. What began as a conference has grown into a series of online events, an annual summit and a growing community [83]. According to the 2023 ITU Report, *United Nations Activities on Artificial Intelligence*, 408 UN projects involving AI connect to one or more Sustainable Development Goals

(SDGs), with approximately 32% (135 projects) specifically addressing SDG 16 (Peace, justice and strong institutions) [83]. This substantial number reflects growing interest in integrating AI tools into development programming and policy. However, the report primarily offers very brief overviews rather than in-depth analyses of implementation methods, effectiveness metrics, or challenges encountered in the field.

Connected to the UN are several other key initiatives advancing practice, research, and collaboration around technology and peace, including AI and emerging technologies. The UN Innovation Network is convening a cross-departmental generative AI working group to foster learning and experimentation throughout the institution. Meanwhile, the Innovation Cell in the Department of Political and Peacebuilding Affairs has piloted the exploration of AI tools for dialogue and engagement in conflict-affected areas [84]. The Economic, Social and Cultural Council (ECOSOCC), established in July 2004, serves as an Advisory Organ of the African Union (AU). Composed of various social and professional groups from AU Member States, ECOSOCC develops online toolkits focused on peace and security issues for civil society organizations, partners, and journalists [85]. The AU is also developing comprehensive policies and best practices through its AI, Security and Peacebuilding working group to enhance effectiveness across the continent and strengthen cooperation with civil society organizations. The Global Peace Tech Hub at the European University Institute is another important initiative. It has partnered with The Gov Lab and the Institute of Social Ethics at the University of Lucerne to create a database of PeaceTech initiatives and conduct-related research [86], [87].

Other efforts include those of the Council on Technology and Social Cohesion, a multi-stakeholder initiative launched in 2023 that brings together diverse actors to explore how technology can foster trust rather than drive polarization [88]. Significant international frameworks have also emerged, including UNESCO's Recommendation on the Ethics of Artificial Intelligence, which provides the first global standard-setting instrument for ethical AI development [89]. Additionally, the United Nations Hub for Human Rights and Digital Technology addresses the human rights implications of digital technologies, including AI [46], [56]. In September 2024, the UN adopted the Global Digital Compact

at the Summit for the Future—a framework promoting responsible digital technology and AI use while addressing potential harms [90].

For over a decade, Build Up, a non-profit organization, has connected peace and technology communities through training, research, and consulting. Since 2014, it has hosted an annual global conference and offered fellowship programs supporting pilot work at the intersection of AI and peacebuilding [91]. The Alliance for Peacebuilding, in partnership with Search for Common Ground, the Toda Peace Institute, and Mercy Corps, has developed a Digital Peacebuilding Community of Practice where members explore tools, research strategies and collaboration [92].

The next section of this Technical Paper will provide an overview of where AI tools are being used and how they can positively contribute to advancing peace and reducing conflict.

SECTION 6. APPLIED USE OF AI TOOLS

Overarching Risks and Limitations in AI Peacebuilding Applications

Before examining the specific uses of AI connected to peacebuilding, it is valuable to review common limitations and risks. The deployment of AI tools in peacebuilding faces several fundamental challenges that span different applications. Privacy and security concerns are paramount when collecting data from vulnerable populations in conflict zones. Key challenges include data security, ethical collection methods, and transparency about how data will be used [93], [94], [95]. The digital divide can exacerbate inequities, with more reliable data typically coming from urban areas while neglecting remote or digitally marginalized communities [29], [96]. Moreover, AI tools can often be used for surveillance, predictive policing or other more nefarious purposes.

Data collection in conflict zones presents persistent difficulties due to reduced media production, access limitations, and potential bias when content comes from limited sources [97], [98], [99]. What is perceived as legitimate data often skews toward quantitative approaches from the Global North, excluding valuable local information [13].

Resource sustainability presents ongoing challenges for AI implementation in peacebuilding. Despite decreasing costs, AI systems require substantial resources—particularly larger systems needing extensive training data, refined algorithms, computational power, and significant human input and analysis [45], [46]. Securing sustainable funding for both human-centered work and AI systems remains difficult.

Moreover, the effectiveness of AI tools depends heavily on human oversight. Proper training in verification, interpretation, and ethical application remains essential [100], [101]. There is also risk that overreliance on technology could diminish valuable human intelligence networks that provide crucial contextual understanding [102], [103].

AI systems face inherent limitations in handling uncertainty in both incoming data streams and predictions [16]. While these models can forecast risks with some accuracy, they struggle to account

for unexpected dynamics and “black swan” events—rare but highly impactful occurrences [104], [105]. This challenge is particularly acute in the current era of exponential data growth. Despite the increasing abundance of data about peace and conflict trends, more information does not automatically translate into improved policy or practice [4], [106]. A fundamental gap persists between information and action.

This implementation gap exists because effective peacebuilding, like any policy action, requires navigating complex political and economic realities. Practitioners face critical obstacles: competing priorities, risk-averse bureaucracies, and outcome uncertainties. While data and AI can help inform decision-making, more data in and of itself cannot bridge the fundamental divide between knowledge and action [106], [107].

Conflict Prediction and Forecasting

Potential

AI tools show promise for conflict prevention work despite current limitations in accuracy and gaps in peacebuilding data. These technologies could enhance forecasting of conflict escalation patterns and inform more effective response strategies.

A fundamental assumption in data-driven peacebuilding is that better prediction systems can enable interventions before conflicts escalate into mass violence [29], [108]. If policymakers could better anticipate where and how violence might escalate, they could deploy targeted diplomatic interventions and policies that could potentially save lives, protect infrastructure, and reduce long-term societal and economic damage. AI-assisted conflict-mapping systems might identify early warning signals—such as increasing dehumanizing rhetoric, declining inter-group contact, rising tensions, isolated violent incidents, unusual weapons movements, and weakening institutional trust—possibly allowing for more timely intervention before conflicts intensify to dangerous levels.

Several significant conflict prediction models have been developed in recent decades, such as the Uppsala Conflict Data Program (UCDP). The most prominent is the UCDP Dataset for tracking armed conflict and organized violence. Maintained for over 40 years, this resource serves researchers,

policymakers, and media organizations focusing on conflict trends, impacts, and theoretical models [16], [109]. Updated annually through various reports, it has primarily functioned as a historical analysis tool rather than an effective predictor of future conflicts [16], [97], [110].

To strengthen conflict prediction, researchers at Uppsala University created the UCDP Candidate Events Dataset, which delivers monthly data updates, in addition to annual ones, enabling analysts to better monitor and identify emerging patterns of violence. Scholar Mihai Croicu identifies three factors driving improved forecasting: a "revolution in data" with village-level detail in conflict zones, advancement beyond "shallow machine learning" to robust models using LLMs and deep neural networks, and better integration of data with theoretical frameworks [16], [111].

The Violence and Early-Warning System (ViEWS), a collaboration between Uppsala University and the Peace Research Institute, Oslo, launched in 2017 and was a significant advancement in this field. Building on the Uppsala datasets, on Armed Conflict Location and Event Data (ACLED), and Varieties of Democracy (V-Dem), the project initially aimed to generate monthly forecasts for conflicts up to 36 months in advance [110], [112]. As the field evolved with new datasets and modeling advancements, the project expanded to reflect a broader focus on both predicting conflicts and their impacts. While these models effectively predict high-intensity violence in known conflict zones like Nigeria and Yemen, they still struggle with long-term forecasting and identifying emerging conflicts in unexpected regions [106].

ACLED has been operating since 2005, collecting information from local media, human rights reports, and various sources globally. The organization specializes in tracking specific event types (including protests, battles, and violence against civilians) while documenting involved actors, violence categories, and fatalities. This data enables ACLED to produce short-term "conflict pulse" alerts [113], [114].

Both the Uppsala Conflict Data Program and ACLED analyze key indicators and metrics to understand conflict trends, with ACLED particularly focusing on categorizing event types and tracking the actors involved, violence typologies, fatalities, and related factors [114]. ACLED's

Conflict Alert System (CAST) forecasts changes in conflict events up to six months in advance, reportedly achieving accuracy rates of up to 95% [106], [113].

Uppsala's models, meanwhile, examine political violence types, conflict history, structural economic factors, and news sources [78]. The ViEWS model forecasts expected battle deaths from political violence, contributing to a better understanding of conflict and peace drivers [112].

One of the trends of AI in 2025 is the growth of AI agents, which can be tasked with doing research, completing tasks, and carrying out actions on behalf of an individual or organization. If reliability improves in terms of output, these agents could make use of all of these complex and nuanced datasets, in addition to social media, news, and other channels to perform real-time semantic analysis, even more accurate alerts, and potentially even policy recommendations.

Risks Specific to Conflict Prediction

There are several key risks to AI-facilitated conflict prediction. First, forecasts can be misused and lead to unintended outcomes. If an AI system predicts a shift in conflict dynamics, parties to the conflict might preemptively escalate violence to secure territorial or political advantages before these anticipated changes occur [110]. Second, conflict forecasts might trigger external interventions, leaving conflicts unresolved without formal peace agreements or reinforcing power imbalances between parties [29]. Third, predictions may raise false hopes of external intervention. Fourth, because forecast accuracy remains limited, there is a risk of delegating complex political decisions to potentially biased AI systems. Finally, security-oriented approaches may label legitimate social movements as threats based on AI training biases, leading to persecution. Examples include China's targeting of Uighurs and the USA Government's use of AI to identify pro-Palestinian student protesters [56], [115].

Real-Time Mapping of Conflict and Humanitarian Contexts

Potential

Satellite imaging, drone surveillance, and crowd mapping have improved conflict monitoring and humanitarian response efforts. These technologies can complement each other when deployed in conflict settings.

- Satellite imagery can help track important developments in hard-to-reach conflict zones, including destruction of infrastructure, movement of fighters and weapons and displacement trends [116].
- Drones can deliver detailed assessments in dangerous or restricted conflict areas.
- Crowd mapping facilitates widespread civilian participation, helping to create near real-time documentation of events such as violence, protests, assessing needs after a human or natural disaster [117], [118].

These technologies have proven particularly valuable in active conflict zones, where traditional monitoring faces severe limitations due to security concerns, access restrictions, and rapidly changing conditions.

The integration of generative AI has dramatically enhanced these conflict-mapping capabilities. What once required extensive manual analysis can now be processed much more rapidly, with AI systems identifying patterns of destruction, population movement, and military activity from visual and geographic data. While human oversight remains essential for accuracy and verification, processing speed has increased exponentially.

Crisis mapping has evolved significantly since its early applications in Kenya (2008) and Haiti (2010) [118]. Initially reliant on trained volunteers manually tagging and categorizing data, the field now leverages machine learning to analyze inputs from social media, satellites, and drones. Organizations like Ushahidi and the Humanitarian OpenStreetMap Team (HOT) have transformed grassroots initiatives into operations that partner with local groups, UN agencies, and governments worldwide [96], [119].

AI integration in conflict mapping is producing significant improvements. With high-quality training data from previous conflicts, AI systems can now identify military assets, damaged infrastructure, and population movements with increasing accuracy and speed [120]. The UN and the International Criminal Court regularly use satellite imagery enhanced by AI analysis to document human rights violations, while drones monitor troop movements in peacekeeping missions [121]. Private companies like Planet have expanded capabilities further, launching satellites that capture thousands of daily images worldwide, which is proving to be invaluable for documenting war damage in Ukraine [122]. Organizations like the United Nations Satellite Centre (UNOSAT) employ these technologies to document cultural heritage sites in conflict zones, creating digital archives that may aid reconstruction efforts in post-conflict settings and provide evidence of war crimes targeting cultural property [123].

Advanced systems like Data Insights for Social and Humanitarian Action (DISHA) now integrate satellite imagery, drone footage, phone data, and crowdsourced reports for humanitarian response. DISHA is led by the UN Global Pulse in partnership with numerous agencies including UNOSAT, the International Committee of the Red Cross (ICRC), the World Food Program, and Google.org [124]. The system is being tested, and a comparison was conducted across nine natural disasters to examine how AI analysis would perform compared to more manual assessments [125]. The early results are promising, showing that damage assessments could be potentially reduced by “a time factor of six” while also expanding the range of assessments [125]. While this finding is relevant for humanitarian relief efforts, DISHA will likely have strong uses and impact in conflict-affected settings.

The Danish Refugee Council (DRC) has pioneered displacement prediction models that enable more accurate forecasting for budgeting and resource allocation [126]. The Data Entry and Exploration Platform (DEEP), developed by the DRC, integrates AI and natural language processing to enhance collaboration among humanitarian actors through improved resource searches and automated reporting [127]. Similarly, the ICRC employs machine learning dashboards to improve staffing and resource allocation for field programming.

Limitations Specific to Mapping

Platforms such as DEEP face persistent sustainability issues, including precarious funding streams and troubling engagement metrics—with a mere 20% of registered users demonstrating active annual participation [128]. Serious questions persist regarding the network's long-term financial viability and operational sustainability, particularly since it was primarily funded by USAID.

Risks Specific to Mapping

AI technologies carry serious "dual-use" concerns as they can be used for both humanitarian mapping and military targeting and surveillance. When AI systems categorize certain identity groups as security threats, members of these groups can face heightened risk of rights violations [94]. Data contributors in conflict zones also face security dangers if their identities or locations are compromised [94].

Digital Inclusion and Citizen Engagement

Potential

AI technologies are contributing to peacebuilding efforts by enhancing dialogue and engagement around contentious issues. Similar to some online dispute resolution systems, AI-powered engagement platforms leverage AI as a process facilitator—helping to identify common ground, analyze complex data, map participant interactions, and support community or policy deliberation [114], [129], [130], [131], [132]. What makes AI particularly valuable in these contexts is its capacity to dramatically scale engagement across diverse groups, reaching across geographic, language, and identity divisions that might otherwise limit participation [73], [133].

In conflict-affected regions, strategic partnerships between the UN and technology companies such as Remesh, have expanded citizen participation in policy development and dialogue initiatives. Governments and civil society groups have frequently used another tool, Polis, a civic engagement platform that has enabled thousands of participants to contribute to public policy discussions in Taiwan and elsewhere. Facilitators can use the algorithms and digital tools to identify areas of

consensus among diverse viewpoints and engage participants in constructive explorations of divisive issues on a scale that would be impossible through human analysis alone [101], [134].

The UN Department of Political and Peacebuilding Affairs (DPPA) has used Remesh in conflict zones including Libya, Yemen, and Syria, engaging up to 1,000 participants per initiative. Each digital consultation was tailored to address specific regional contexts, conflict dynamics, and demographic considerations, enabling broader representation in dialogue, policy and/or peace processes [133], [135], [136]. As explained by Irwin et al., "The Libyan Digital Dialogues addressed the impact of the civil war and ceasefire, domestic militias and foreign fighters, economic issues—including a fair distribution of oil revenues—as well as concerns around human rights and future elections" [p.3, 127]. Some sessions were moderated by the Acting Special Representative to Libya, sessions that generated questions for Government of National Unity candidates, who responded on broadcast television and over social media.

In Yemen, the UN organized public dialogues, with AI support, to create opportunities for citizen input on national priorities. A June 2020 dialogue engaged 500 participants over three hours. Daanish Masood of DPPA's Innovation Cell notes the approach "represents a new way of doing business that can make ongoing political and peace processes far more inclusive" [p.2, 125]. A more recent example is the Alliance for Middle East Peace's partnership with Remesh. In 2024, they organized peacebuilding dialogues both within and between Israeli and Palestinian peacebuilding communities [137], [138]. Remesh was used to frame questions, analyze data, and highlight commonalities to explore potential opportunities for peacebuilding. Ultimately, organizers hope to expand the pool of participants and that the results will inform policy and advocacy work.

In Sudan, Crisis Management Initiative–Martti Ahtisaari Peace Foundation (CMI) used Remesh to engage over 1,000 participants in July 2023, focusing on women's groups, networks and alliances and youth and resistance committees to obtain insights on priorities and representation [139]. CMI invested significantly in participant recruitment and consent, sharing its findings with the participants and leveraging the outcomes with its partners [139].

Another tool that has been used in numerous settings is Talk to the City which was developed in 2023 to support large-scale dialogues, leveraging the power of customized LLMs to analyze discussions, map out common ground and show areas of overlap and consensus [140]. CMI used WhatsApp chat bots with youth in Yemen and then used Talk to the City to map out results for sensemaking.[4]

Limitations Specific to Digital Inclusion

There are several key limitations specific to dialogue and civic engagement work that leverages AI tools. In conflict-affected settings in particular, it can be difficult to determine the impact of such processes on advancing peace or improving interactions. Almost all the literature is descriptive rather than evaluative, and lacks metrics and an assessment of effectiveness. Moreover, unless the process is longer term, its impact may be limited.

Access remains a significant consideration. Despite efforts to accommodate low bandwidth environments, connectivity issues can make it challenging to include participants from resource-poor environments. There is also the issue of the cost for those needing SMS or internet connectivity to participate in the processes, as well as the potential costs of running the software for providers. While some digital AI software companies may offer their software for free or at a discount for specific uses by public or nonprofit organizations, this may not always be the case [134].

Pro-Social Engagement through AI

Potential

AI can serve as a valuable tool within social media environments. First, it can moderate content and mitigate potential harms such as hate speech, echo chambers, and polarization, including detecting misinformation or disinformation. Second, it can foster more pro-social engagement and connection

[4] Michelle Giovanardi, “Information on the CMI project”, personal communication with Craig Zelizer, video call, March 7, 2025.

between users. Despite the initial vision of social media as a great connector, platforms have caused and continue to inflict significant harm. As Fredrick Ogenga notes, “Social media is a double-edged sword which can be used both for peace and for conflict”[141].

Mitigating Harms and Content Moderation

Prior to the advent of generative AI, machine learning algorithms were widely deployed by social media companies primarily to foster greater online engagement. This questionable business practice has been extensively evaluated, revealing a troubling pattern across many platforms: content triggering outrage and divisiveness consistently generates more engagement and has been prioritized to generate more revenue by many leading platforms [142], [143], [144], [145].

An increasingly prominent approach in social media involves using machine learning and AI tools to track posts across X, TikTok, Facebook, WhatsApp, and other platforms. Though human moderation remains essential, algorithmic scraping and analysis can enable the tracking of hate speech and extremist and offensive content as well as the identification of disinformation and misinformation. Such analysis is very important for monitoring conflict trends, escalation, or peacefulness in specific areas. For example, Build Up has developed Phoenix, an open source tool that analyzes public posts on social media channels to identify potential areas for peacebuilding engagement [146]. There are also efforts to go beyond just using language to detect misinformation or hate speech and to leverage AI tools and human analyses to look at behavioral signals, such as friending certain people or outlets, engaging in content, spreading content and related items [28], [73], [129].

One example that combines old technology and new is the United Wave project, originally deployed by UN Global Pulse and now run by the UN Office of Information Technology [147]. The project uses AI to collect, transcribe, and analyze radio broadcasts and is being deployed in numerous peacekeeping missions and other UN agencies.[5] Radio remains a powerful means of communication

[5] Joseph Aylett-Bullock, “Email About AI”, personal communication, email to Craig Zelizer, May 4, 2025.

dissemination in many parts of the world, and the ability to monitor for conflict escalation, trends, and disinformation using AI provides near real-time data that is invaluable for planning and programming.[6]

Fostering Pro-Social Engagement

Various peaceful activities have been organized through social media to support social movements, civic engagement, skills building and advocacy [108], [148], [149]. While social media has contributed to the polarization and escalation of violence in Myanmar, Kenya, Colombia, the United States, and the Philippines, groups have also leveraged Facebook, WhatsApp, and other platforms to mobilize for peace [148], [149], [150]. In Kenya, the Center for Media, Democracy, Peace and Security at Rongo University along with other local partners, in collaboration with Build Up created Maskani (or "home platform"), a local digital peacebuilding infrastructure that employed WhatsApp and Facebook to address ethnic political polarization both online and offline during the 2022 presidential elections [148]. While the project didn't incorporate AI elements, it is a strong example of building positive peace digital communities by using new technologies for new channels of engagement.

A small but growing movement is shifting the focus toward AI projects that prioritize "pro-social" interactions rather than merely avoiding harm. Unlike the defensive emphasis on harm mitigation and safety in mainstream AI, pro-social technology deliberately fosters meaningful human connections and bridge-building across divides [15], [73], [151], [152]. This approach requires reimagining the process of building AI tools by integrating ethics and community-building principles from the ground up—influencing everything from initial design to deployment and governance. As Lisa Schirch observed at the Designing Tech for Social Cohesion conference, "The road to Hell is

[6] This project is adapted from early radio monitoring work around peacekeeping and health initiatives. For more information, see: UN Foundation, "When Old Technology Meets New: How UN Global Pulse is Using Radio and AI to Leave No Voice Behind", https://unfoundation.org/blog/post/when-old-technology-meets-new-how-un-global-pulse-is-using-radio-and-ai-to-leave-no-voice-behind/

paved with code, and if we want to pave a road out of Hell we need to think about how we pave a different road with different code" [145].

The team behind Perspective API has recently shifted focus toward exploring how to better incentivize constructive dialogue across diverse perspectives. Instead of sorting conversations chronologically, they reported that using AI to sort conversations into meaningful categories—reasoning, personal stories, and curiosity—resulted in readers finding these exchanges "not only less hostile but also more informative, respectful, trustworthy, and interesting" [151].

Limitations Specific to Social Media

Content moderation systems face significant limitations unique to social media contexts. These systems can misclassify content—falsely labeling benign material as harmful or failing to flag genuinely problematic speech [153]. They struggle to accurately process linguistic nuances and cultural contexts. There is also a move by X and Facebook to incorporate community notes as a means of moderation which places the responsibility more on the users and community. Community moderation faces several challenges, including increased risk of disinformation spread, moderators' limited expertise and time constraints, tendency toward oversimplified judgments, and various other operational issues [154]. The rapid proliferation of AI-generated "slop content"—low-quality text riddled with inaccuracies, poor grammar, and fabricated information—presents a mounting challenge that current content moderation systems appear ill-equipped to handle [155]. Also defining hate speech or toxic content can be inherently subjective and highly context-dependent.

Although there is significant potential for pro-social engagement tools to have a positive impact, their deployment remains relatively limited. Moreover, these types of tools encounter substantial challenges regarding sustainable funding models and struggle to develop business approaches that can both support ongoing operations and facilitate broader implementation across diverse contexts.

Risks Specific to Social Media

As social media platforms increasingly rely on algorithms to shape user engagement, several concerning trends are emerging. First, as AI is assuming a greater role in content generation and interaction, human agency is diminishing. Digital spaces risk becoming environments where algorithms primarily engage with other algorithms, fundamentally altering what constitutes meaningful human connection and pro-social interaction. This is particularly critical where so much of peacebuilding work has relied on building pathways to improved relationships and understanding between groups in conflict. As Goldberg et al. comment, “As AI technologies evolve, it will be even easier for small groups to dominate conversations or create a majority illusion, leveraging ever more realistic bots or agentic AI, raising urgent questions about how to preserve the integrity of the public square” [p. 27, 73].

Second, pro-social or harm mitigation tools may have an artificial bias toward superficial positivity. These systems might suppress productive tension or disagreement that, if thoughtfully navigated, could foster deeper relationship-building and authentic understanding [74].

AI in Mediation and Negotiation Process

Potential

Online dispute resolution (ODR) has evolved from making conflict resolution more accessible and cost-effective to incorporating sophisticated AI capabilities. Hibah Alessa identifies two promising roles for AI in mediation: a support role that enhances process efficiency, and a substitutive role where AI directly facilitates or contributes to decision-making in place of a human facilitator [132].

AI shows significant potential for streamlining dispute resolution by automating administrative tasks like notetaking, intake processes, and case management. More advanced applications offer functions such as mapping complex options, analyzing large datasets, facilitating common ground, engaging with distributed participants, and generating insights that human mediators might miss.

Major technology platforms demonstrate AI’s scalability in dispute resolution. By 2019 eBay was resolving 60 million cases annually with its AI tools. More recently, Mercado Libre’s Verdi AI platform

is handling 10% “of the e-commerce company’s customer service disputes on one of its major sites” [156], [157]. Although the use of such systems is a long way off in peacebuilding, these types of systems could provide valuable data-related insights and help inform peacebuilding processes in the future.

Recent research suggests that AI may effectively help bridge social divides. A UK study with several hundred participants found that LLMs sometimes outperformed human mediators as “caucus mediators” when framing shared interests between groups with opposing viewpoints on divisive policy issues [158]. These emerging applications show promise for real-world dialogue.

AI agents working within human-defined parameters can negotiate disputes between parties. For instance, in commercial conflicts, opposing sides might delegate negotiation to AI representatives programmed with specific conflict-related approaches (collaborative or competitive) to generate resolution options. While AI agents are unlikely to replace human negotiators in high-stakes international conflicts, they can support scenario planning. The US military has tested various LLMs with classified data to simulate potential responses to Indo–Pacific regional conflicts and develop strategic forecasts [159], [160].

Peacebuilders will likely be able to leverage AI systems to better explore negotiation scenarios and outcomes and to test different strategies. Increasingly sophisticated negotiation bots will likely be able to provide real-time insights during negotiation processes regarding tactics, strategies, and options. Researchers and practitioners are starting to explore how LLMs and AI tools can assist in diplomatic negotiations and preparations. For example, the German Foreign Office in partnership with the AI company, Omdena, ran a proof of concept to see how integrating trained LLMs would impact policy officers [161]. This project sought to aggregate knowledge by gathering data from relevant documents, building an LLM interface, training officers on the tool, and exploring the results. In particular, the group reported up to 70% time savings in searching for the documents needed for negotiation preparation and research [161].

Limitations Specific to Mediation & Negotiation

Although AI can often effectively handle structured disputes that are common in e-commerce, complex peacebuilding processes in conflict zones require nuanced judgments that are difficult to automate. AI-first negotiation or mediation works well for addressing routine customer problems, where much of the process follows predictable patterns. However, in situations involving threats of violence or escalating tensions, deploying AI directly as a third-party mediator could produce dangerous outcomes. In complex conflicts, AI tools are best used in an assistive rather than a directive capacity—providing relevant data, processing insights, or helping human mediators plan strategic interventions.

Risks Specific to Mediation/Negotiation

Peacebuilding practitioners are concerned about the potential weakening of human facilitation skills. First, as reliance on AI tools for pro-social engagement increases, practitioners may become "process lazy", losing the nuanced interpersonal abilities that are essential for effective peacebuilding [131], [162]. Second, they risk developing dependencies that may erode human decision-making capacity during crisis responses—particularly dangerous in volatile contexts [120], [162]. The uneven distribution of AI capabilities threatens to widen technological disparities between well-resourced international organizations and local peacebuilding actors.

SECTION 7. BUILDING A HEALTHY INFORMATION ENVIRONMENT

AI can contribute to peacebuilding research, practice, and policy when ethically deployed. The recent history of such dual-use technologies, however, reveals that sophisticated machine learning systems can be used for surveillance, spreading disinformation, violating rights, and deepening divisions in conflict zones. A healthy information environment is foundational to leveraging AI for peacebuilding. Without it, even the most sophisticated AI tools may exacerbate rather than prevent conflict.

The key elements of such an environment include diversity of voices, with media ecosystems supporting varied content creation and constructive public policy oversight [7], [136]. Citizen engagement is equally important, alongside strong accountability mechanisms [136].

Digital literacy is crucial, as citizens must be equipped to find trusted information, identify misinformation and disinformation, protect their privacy, and understand AI capabilities and limitations [129], [136]. Minimal manipulation by information operations, particularly from domestic sources, helps maintain information integrity, while societal prioritization of reliable information supports truth-seeking behaviors [13], [163].

AI may be especially useful in peacebuilding and making information environments resilient when new tools are accompanied by human capacity development and effective policy frameworks. In developing new AI applications, information verification tools can detect and flag misinformation in conflict contexts. Information accessibility through real-time interpretation and improved citizen services can bridge communication divides [5]. AI-assisted content moderation can reduce harmful content while preserving diverse perspectives. It is important to note that human oversight is key given the hallucinations, inaccuracies and biases that still permeate AI models.

Human capacity development can encompass comprehensive digital literacy training aimed at empowering citizens to critically evaluate information, understand the nuanced implications of AI in conflict contexts, and build user proficiency, including practical skills for effective prompting.

Customized learning experiences, carefully tailored to address the specific needs and contexts of individuals, sectors, and organizations, offer considerable potential to enhance outcomes and effectiveness [159]. Furthermore, targeted professional development opportunities for journalists, civil society actors and policymakers may become particularly important to information integrity and democratic resilience.

Effective policymaking can be strengthened through meaningful civic participation that can be enriched with AI tools. Policymakers can facilitate greater public dialogue, clearly identify shared interests, and seek direct citizen input with several of the platforms discussed above. Crucially, addressing the challenges posed by AI requires stronger partnerships and collaboration across sectors—including government, civil society, academia, peacebuilding organizations, and technology companies—to ensure both the responsible development of AI and effective accountability measures.

Future Areas of Opportunity

As AI advances, it will play an increasingly significant role in peacebuilding. Key applications beyond those listed in this Technical Paper include conflict resolution training and upskilling, supporting peacebuilding ecosystem learning, knowledge management and effectiveness, skills development, and facilitating engagement around mental health [39], [120], [133].

Peacebuilding Infrastructure and Effectiveness

A fundamental challenge facing peacebuilding practitioners is identifying and prioritizing interventions that yield the greatest impact in conflict-affected regions—something AI may also assist with. Given constrained resources and complex contexts, effectively selecting and sequencing peacebuilding activities can be difficult and resource-intensive [77], [164]. As evidence about peacebuilding effectiveness continues to accumulate, practitioners must navigate trade-offs between short-term needs and long-term outcomes, balancing immediate responses against sustainable strategies. AI may also help diplomats and negotiators evaluate such trade-offs.

Artificial intelligence offers promising tools to help address these complex challenges. AI systems could systematically analyze extensive datasets—including programming reports, evaluations, stakeholder analyses, and impact assessments—to recommend evidence-based and context-sensitive interventions. By leveraging advanced modeling, scenario planning, and comparative cost-effectiveness analyses, these tools may significantly streamline decision-making processes, enabling practitioners to identify optimal interventions, efficiently manage limited funding, and strategically sequence activities for maximum impact.

Furthermore, AI platforms have the potential to improve collaboration and knowledge sharing among peacebuilding stakeholders. Initiatives such as the DEEP platform illustrate how AI-driven collaboration can facilitate exchange of best practices and uncover potential partnerships. AI can also strengthen knowledge management by allowing organizations to securely organize and rapidly access their internal documentation, evaluations, and technical reports. For example, the Inter-American Development Bank's recent implementation of Seek AI illustrates AI's potential for enhancing knowledge accessibility and public engagement, as it allows users to interact directly with thousands of peer-reviewed resources using generative AI tools [165].

While AI holds considerable promise, it should complement rather than replace human expertise and local partnerships. Peacebuilding is inherently context-specific, shaped by political, cultural, and social dynamics that are often difficult to fully capture as data. Technically sound interventions risk failing to gain traction unless combined with human oversight and genuine community engagement. Thus, human decision-making and local collaboration need to remain central elements of effective peacebuilding, with AI serving as a powerful supportive tool.

Skills Development

AI can help individuals, communities, and organizations develop conflict resolution skills through personalized training. There is tremendous potential to use AI tools to provide training for diplomats, peacebuilders, development workers, and others engaged in negotiation and peacebuilding work. Organizations could create customized learning processes for staff working in

conflict zones by leveraging both open and proprietary data, reducing development time significantly.

For example, the staff of international organizations who are deployed to conflict-affected regions could receive AI-assisted cultural training, be trained to understand conflict contexts, learn how to conduct stakeholder analysis, and improve their negotiation skills. Like Khan Academy's Khanmigo platform, AI assistants could provide direct support for peacebuilding education. A promising initiative lies in developing smaller, customized LLMs for internal institutional training, prioritizing data relevant to the organization's or sector's needs.

It is now becoming possible to leverage AI to create customized learning programs based on an organization or individual's particular goals. Companies like Arist or Unschooler are helping companies and educational institutions build high-quality learning pathways for their communities and the potential to do so in the peacebuilding field is enormous [166], [167].

Mental Health and Reconciliation

A key challenge in many conflict-impacted areas is that long-term displacement, violence and suffering can also lead to large scale stress, potential post-traumatic stress disorder and other mental health challenges. Currently, AI tools are being deployed for wellbeing purposes in numerous settings. Several organizations are building trauma-sensitive chatbots to provide support or information services to people impacted by domestic violence [168]. In conflict zones, where mental health support resources are often severely limited, carefully designed AI interventions could potentially fill some gaps in the information environment for those unable to access traditional support services.

It is crucial to highlight the ethical concerns and challenges arising from shifting beyond using AI tools as service support to deploying chatbots directly with individuals in conflict-affected settings [169]. While these tools risk causing significant harm if misused, recent research shows promising preliminary results [170]. Under rigorous ethical guidelines—including transparency about AI limitations, cultural sensitivity protocols, human oversight mechanisms, and clear escalation

pathways—AI chatbots could provide targeted support to conflict-affected populations when properly integrated into existing humanitarian frameworks [39]. Rather than immediately pursuing AI-assisted therapy, a more prudent approach might involve AI coaches or resources that complement individuals' existing coping processes [39]. Additionally, AI-enhanced virtual and augmented reality applications may offer meaningful contributions to this field [171].

SECTION 8. RECOMMENDATIONS FOR THE USE OF AI IN PEACEBUILDING

As AI rapidly evolves, numerous opportunities emerge to integrate AI tools into efforts aimed at fostering peace and preventing conflict. To maximize AI's potential for positive impact in peacebuilding and related fields, the following four key recommendations outline essential considerations for safely leveraging AI across sectors. Subsequent recommendations tailored specifically to sectoral stakeholders are also provided.

Recommendations for Design and Application

Rights-Based, Conflict-Sensitive Design

Designers and practitioners must apply “do no harm” principles throughout the AI development lifecycle for peacebuilding projects and prioritize human rights and consent. They must embed ethical frameworks from design through deployment that acknowledge power dynamics in fragile contexts and recognize technology's political dimensions. Ultimately, the organizations that use AI in peacebuilding should create certification systems that recognize technologies designed primarily for constructive social outcomes.

Inclusive Development and Digital Access

The parties involved in peacebuilding efforts need to close the digital divide by increasing the resources available to Global Majority countries, support AI applications in diverse languages, and facilitate regional AI development. New peacebuilding projects must be co-created with affected communities from inception, so that diverse perspectives shape technology at the design phase, rather than treating inclusion as an afterthought. Smaller LLMs and AI trained on relevant databases are more likely to serve local peacebuilding needs.

Continuous Risk Assessment and Knowledge Sharing

Peacebuilding policymakers and researchers should regularly evaluate potential unintended consequences with special attention to dual-use risks. They must document both successes and failures, making the data accessible to a wider community.

Capacity Building Beyond Technical Skills

Civil society, government and the private sector should help expand digital literacy on AI's potential, limitations, and risks across different stakeholder groups, particularly when AI is being utilized in peacebuilding processes. The donor community should encourage cross-sector collaboration that connects peacebuilders with technologists, balancing technical expertise with contextual knowledge while avoiding techno-solutionist approaches.

Sector-Specific Recommendations

Policymakers and Public Officials

The people who support peacebuilding initiatives on behalf of government play a critical role in establishing inclusive regulations and oversight mechanisms that protect human rights while enabling innovation in peace technology. Policymakers can expand funding for AI-enabled initiatives aimed at conflict prevention and resolution and ensure that peacebuilding programs and tools are developed collaboratively, incorporating meaningful input from communities directly affected by conflict. Policymakers can also support research into AI tools that foster citizen engagement and pro-social behavior, as well as tools designed to track and mitigate AI-related harms. International cooperation on AI governance tailored to the specific needs of fragile states is crucial, as is fostering greater collaboration between peace practitioners and technology policymakers through targeted funding and fellowship initiatives.

Private Sector

Companies have an important responsibility to partner directly with peacebuilding organizations, ensuring diverse teams guide product development and create region-specific AI tools appropriate for local languages and contexts. Peacebuilding is an important use-case for new AI tools, and the

private sector should design for the kinds of applications discussed above. They must commit to fair labor practices throughout AI supply chains and reduce environmental impacts. Establishing talent exchanges and partnerships with peace-focused organizations can strengthen cross-sector cooperation and understanding. Additionally, technology businesses can help educate society about risks associated with AI technologies, including misinformation, disinformation, and deepfakes, and share approaches to effectively mitigate these challenges.

Civil Society

Civil society organizations need staff with technology and peacebuilding expertise, ensuring affected communities can actively participate in designing and implementing AI solutions. Broad AI governance, and focused implementation guidelines, must be rooted in human rights and participatory principles, so that civil society groups apply AI safely in conflict-sensitive environments. Encouraging innovative partnerships to address AI harms and offering fellowships and support for technologists and peacebuilders to exchange skills can further strengthen capacity. Civil society can also involve affected communities in monitoring the real-world impacts of AI and leverage existing tools to enhance knowledge management and upskilling.

Donor and Philanthropic Community

Funders providing development assistance can increase support for innovation and research examining AI's potential and limitations in peacebuilding contexts. They should actively invest in initiatives addressing biases and technology access inequities, especially supporting Global Majority organizations developing locally relevant solutions. Longer-term engagements, as opposed to short-term pilot projects, enable adaptation and sustained impact. Development assistance, whether from private philanthropy or government development funds, can also support fellowships and capacity-building to strengthen expertise at the intersection of AI and peacebuilding, while documenting effective practices and learning from diverse contexts.

Academia

Academic institutions can support collaborative research with stakeholders in real-word settings, translating all of these principles into practical solutions. Producing curricula that build technological proficiency alongside the insights of peace and conflict studies is critical to develop the next generation of peacebuilding policymakers, researchers and practitioners. Studying AI applications in peacebuilding contexts and creating interdisciplinary research hubs focused on ethical technology solutions for conflict transformation are valuable ways forward. Connecting students with AI experience and training to peace-focused organizations through fellowships and practical opportunities can further enhance real-world relevance and impact.

SECTION 9. CONCLUSIONS

It is clear that as AI becomes increasingly embedded across our social institutions and our civic life, its impact will grow substantially. As explored in this Technical Paper, AI is having positive effects in peacebuilding, development, and humanitarian sectors through improved conflict analysis and mapping, citizen engagement, and logistics. When rooted in ethics, human rights, user concerns, and privacy, AI can dramatically accelerate tasks such as crisis mapping and data analysis, enabling rapid responses by the peacebuilding community.

AI is a useful tool in its current capacity. However, it is dangerous to assume that merely gathering and analyzing enough data will lead to a more peaceful world. Assuming this overlooks the complex nature of the information environment that generates all the data upon which AI is based. Successful peacebuilding is usually dependent on political commitment and financial resources invested in both the information environment and the peacebuilding processes. Delegating complex decisions to machines built within existing systems—without proper ethical frameworks and safeguards—is likely to cause substantial harm. As Claire Wilmot notes in her recent article on AI and peace, "Political problems can have technical challenges, where AI tools could help. But war is fundamentally a problem of power and politics" [172].

As emphasized in this Technical Paper, AI itself is merely a tool with the potential for both positive and negative applications. It is critical to assess its potential, its current uses, and the existing and future risks of using it. Distinguishing AI hype from practical applications requires rigorous research and thoughtful policy development. While AI integration into warfare, autonomous weapons, disinformation campaigns, and threats to civil liberties raise legitimate concerns, a rights-centered framework could advance positive, meaningful applications of AI to peacebuilding efforts.

REFERENCES

[1] K. Schwab, “The Fourth Industrial Revolution: what it means, how to respond,” World Economic Forum. [Online]. Available: https://www.weforum.org/stories/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

[2] V. T. Razakamaharavo, “Implications of emerging technologies on peace and security in Africa,” ACCORD, Aug. 2021. [Online]. Available: https://www.accord.org.za/conflict-trends/implications-of-emerging-technologies-on-peace-and-security-in-africa/

[3] F. Greenwood, “Data colonialism, surveillance capitalism and drones,” in *Mapping Crisis: Participation, Datafication and Humanitarianism in the Age of Digital Mapping*, D. Specht, Ed., University of London Press, pp. 75–92.

[4] R. Mac Ginty and P. Firchow, “The data myth: interrogating the evidence base for evidence-based peacebuilding,” *Data & Policy*, vol. 6, p. e80, 2024, doi: 10.1017/dap.2024.80.

[5] M. Giovanardi, “AI for peace: mitigating the risks and enhancing opportunities,” *Data and Policy*, vol. 6, p. e41, Oct. 2024, doi: 10.1017/dap.2024.37.

[6] A. Wanless and J. Shapiro, “A CERN model for studying the information environment,” Carnegie Endowment for International Peace, Nov. 2022. [Online]. Available: https://carnegieendowment.org/research/2022/11/a-cern-model-for-studying-the-information-environment?lang=en

[7] “Public Information Environment, Trends in the Global Information Environment: 2024 Expert Survey Results,” Nov. 2024. [Online]. Available: https://www.ipie.info/research/sr2024-2

[8] McCarthy, J., Minsky, M.L., Rochester, N., and C.E. Shannon, “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.” Aug. 31, 1955. [Online]. Available: http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

[9] M. Murgia, *Code Dependent: Living in the Shadow of AI.* Henry Holt and Co., 2024.

[10] M. Reichstein *et al.*, “Early warning of complex climate risk with integrated artificial intelligence,” *Nature Communications*, vol. 16, no. 1, 2564, Mar. 2025, doi: 10.1038/s41467-025-57640-w.

[11] N. Ramanathan and J. Fruchterman, “Gather, share, build.,” *Stanford Social Innovation Review*, [Online]. Available: https://ssir.org/articles/entry/ai-for-social-good-data

[12] R. Abbott and B. S. Elliott, “Putting the artificial intelligence in alternative dispute resolution: How AI rules will become ADR rules,” *Amicus Curiae*, vol. 4, no. 3, pp. 685–706, Jun. 2023, doi: 10.14296/ac.v4i3.5627.

[13] R. Mansell *et al.*, “Information Ecosystems and Troubled Democracy: A Global Synthesis of the State of Knowledge on News Media, AI and Data Governance,” International Observatory on Information and Democracy, Paris, Dec. 2024. [Online]. Available: https://observatory.informationdemocracy.org/report/information-ecosystem-and-troubled-democracy/

[14] “Technology and the Future of Online Dispute Resolution: Platforms for Consumer Protection Agencies,” United Nations Conference on Trade and Development, Oct. 2023. [Online]. Available: https://unctad.org/system/files/official-document/tcsditcinf2023d5_en.pdf

[15] A. Konya *et al.*, “Deliberative technology for alignment,” Dec. 06, 2023, *arXiv*. doi: 10.48550/ARXIV.2312.03893.

[16] M. Croicu, “Forecasting Battles: New machine learning methods for predicting armed conflict,” PhD Dissertation, Uppsala University, 2025. [Online]. Available: https://uu.diva-portal.org/smash/get/diva2:1920778/FULLTEXT01.pdf

[17] M. F. Cuéllar *et al.*, “Shaping AI’s Impact on Billions of Lives,” Dec. 11, 2024, *arXiv*. doi: 10.48550/ARXIV.2412.02730.

[18] V. Linardos, M. Drakaki, P. Tzionas, and Y. Karnavas, “Machine learning in disaster management: Recent developments in methods and applications,” *Machine Learning and Knowledge Extraction*, vol. 4, no. 2, pp. 446–473, May 2022, doi: 10.3390/make4020020.

[19] M. Reichstein *et al.*, “Early warning of complex climate risk with integrated artificial intelligence,” *Nat Commun*, vol. 16, no. 1, p. 2564, Mar. 2025, doi: 10.1038/s41467-025-57640-w.

[20] F. M. Simon, S. Altay, S. Altay, and H. Mercier, “Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown,” *Harvard Kennedy Misinformation Review*, vol. 4, no. 5, Oct. 2023, [Online].

Available: https://misinforeview.hks.harvard.edu/article/misinformation-reloaded-fears-about-the-impact-of-generative-ai-on-misinformation-are-overblown/
[21] M. Vaccaro, M. Caoson, H. Ju, S. Aral, and J. R. Curhan, “Advancing AI negotiations: New theory and evidence from a large-scale autonomous negotiations competition,” Mar. 09, 2025, *arXIV*. doi: https://doi.org/10.48550/arXiv.2503.06416.
[22] J. Wei *et al.*, “Chain-of-thought prompting elicits reasoning in large language models,” Jan. 10, 2023, *arXiv*. doi: 10.48550/ARXIV.2201.11903.
[23] I. Belcic and C. Stryker, “What is supervised learning?,” IBM. [Online]. Available: https://www.ibm.com/think/topics/supervised-learning
[24] E. Robinson, D. Egel, and G. Bailey, *Machine Learning for Operational Decisionmaking in Competition and Conflict: A Demonstration Using the Conflict in Eastern Ukraine*. RAND Corporation, 2023. [Online]. Available: https://openlibrary.org/books/OL49776824M/Machine_Learning_for_Operational_Decisionmaking_in_Competition_and_Conflict
[25] M. Chen, “What is unsupervised learning?,” Oracle. [Online]. Available: https://www.oracle.com/artificial-intelligence/machine-learning/unsupervised-learning/
[26] C. Eden and D. M. Kilgour, Eds., “Methods to analyze negotiation processes,” in *Handbook of Group Decision and Negotiation*, Second Edition., Springer Nature, 2021, pp. 40–58. doi: 10.1007/978-3-030-49629-6.
[27] W. J. Yeo, T. Ferdinan, P. Kazienko, R. Satapathy, and E. Cambria, “Self-training large language models through knowledge detection,” Nov. 2024, doi: 10.48550/ARXIV.2406.11275.
[28] J. Hawke, H. Puig Larrauri, A. Sutjahjo, and B. Cerigo, “Understanding to intervene: The codesign of text classifiers with peace practitioners,” *Data and Policy*, vol. 6, 2024, doi: 10.1017/dap.2024.44.
[29] H. Mueller and C. Rauh, “The hard problem of prediction for conflict prevention,” *Journal of the European Economic Association*, vol. 20, no. 6, pp. 2440–2467, Dec. 2022, doi: 10.1093/jeea/jvac025.
[30] B. Panic and P. Arthur, *AI for Peace.* Boca Raton, FL: CRC Press, 2024. doi: 10.1201/9781003359982.
[31] M. Murphy, E. Sharpe, and K. Huang, “The promise of machine learning in violent conflict forecasting,” *Data & Policy*, vol. 6, no. 35, Aug. 2024, doi: 10.1017/dap.2024.27.
[32] A. A. Casilli, P. Tubaro, M. Cornet, C. L. Ludec, J. Torres-Cierpe, and M. V. Braz, “Global inequalities in the production of artificial intelligence: A four-country study on data work,” Oct. 18, 2024, *arXiv*. doi: 10.48550/arXiv.2410.14230.
[33] N. Rowe, “Millions of workers are training AI models for pennies,” *Wired*, Oct. 16, 2023. [Online]. Available: https://www.wired.com/story/millions-of-workers-are-training-ai-models-for-pennies/
[34] R. Hoffman and with GPT-4, *Impromptu: Amplifying Our Humanity Through AI*. Dallepedia LLC, 2023.
[35] “Hallucination Leaderboard,” GitHub Repository. [Online]. Available: https://github.com/vectara/hallucination-leaderboard
[36] W. Knight, “When will AI outsmart humans? Here’s what the expert surveys say,” *Time*, Mar. 30, 2023. [Online]. Available: https://time.com/6556168/when-ai-outsmart-humans/
[37] A. Aguirre, “Keep the future human: Why and how we should close the gates to AGI and superintelligence, and what we should build instead,” Mar. 07, 2025, *arXiv*. doi: 10.48550/arXiv.2311.09452.
[38] F. Federspiel, R. Mitchell, A. Asokan, C. Umana, and D. McCoy, “Threats by artificial intelligence to human health and human existence,” *BMJ Global Health*, vol. 8, no. 5, p. e010435, May 2023, doi: 10.1136/bmjgh-2022-010435.
[39] E. Tauchnitz, “‘Doing Peace’: Conceptualizing relational peace through interactions and networks in a digitalized world",” *Data and Policy.*, vol. 7, no. 2025, pp. 1–21, Apr. 2025.
[40] YouTube. *Jon Stewart & Maria, Ressa On the US’s authoritarian slide*, (Mar. 07, 2025). [Online Video]. Available: https://www.youtube.com/watch?v=jsHoX9ZpA_M
[41] E. Tauchnitz, “Enhancing Global Governance to Safeguard Peace in the Digital Age,” Global Solutions, 2023. [Online]. Available: https://www.global-solutions-initiative.org/publication/enhancing-global-governance-to-safeguard-peace-in-the-digital-age/
[42] “Recommendation on the Ethics of Artificial Intelligence,” UNESCO, 2022. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137

[43] Chief Executives Board for Coordination, “Principles for the ethical use of artificial intelligence in the United Nations System Summary of Deliberations,” United Nations, Oct. 2022. [Online]. Available: https://unsceb.org/sites/default/files/2023-03/CEB_2022_2_Add.1%20%28AI%20ethics%20principles%29.pdf
[44] A. T. Hirblinger, “Building a peace we don’t know? The power of subjunctive technologies in digital peacebuilding,” *Peacebuilding*, vol. 11, no. 2, pp. 113–135, Dec. 2022, doi: 10.1080/21647259.2022.2154957.
[45] United Nations Development Program and G7 Italia, “AI Hub for Sustainable Development: Strengthening Local AI Ecosystems through Collective Action. The G7 Leaders’ Summit in June 2024.” Jul. 17, 2024. [Online]. Available: https://www.undp.org/publications/ai-hub-sustainable-development-strengthening-local-ai-ecosystems-through-collective-action
[46] A. Kriebitz, C. C. Corrigan, A. Pevkur, and A. Horzyk, “Promoting and Advancing Human Rights in Global AI Ecosystems: The Need for a Comprehensive Framework under International Law.,” Ludwig Maximilian University of Munich, the AI Ethics Lab at Rutgers University, the Responsible AI Network–Africa (RAIN-Africa), and the Globethics Foundation., Feb. 2025. [Online]. Available: https://aiethicslab.rutgers.edu/publications/promoting-and-advancing-human-rights-in-global-ai-ecosystems/
[47] A. Kawakami, A. Coston, H. Zhu, H. Heidari, and K. Holstein, “The Situate AI Guidebook: Co-designing a toolkit to support multi-stakeholder early-stage deliberations around public sector AI proposals,” presented at the CHI Conference on Human Factors in Computing Systems (CHI ’24), May 2024, pp. 1–22. doi: 10.1145/3613904.3642849.
[48] A. Kaspersen and W. Wallach, “Why are we failing at the ethics of AI?,” Carnegie Council for Ethics in International Affairs, Nov. 2021. [Online]. Available: https://www.carnegiecouncil.org/media/article/why-are-we-failing-at-the-ethics-of-ai
[49] E. Tauchnitz and S. Ahmed, “Framing ethical e-governance: A plaidoyer for a human-rights based digital democracy approach,” presented at the 17th International Conference on e-Democracy and Open Government (CeDEM 2024), ACM, 2024, pp. 22–27. Accessed: Mar. 30, 2025. [Online]. Available: https://www.dataeconomypolicyhub.org/items/framing-ethical-e-governance:-a-plaidoyer-for-a-human-rights-based-digital-democracy-approach
[50] E. Fournier-Tombs, E. Albrecht, and R. Brubaker, “Disinformation and Peacebuilding in Sub-Saharan Africa: Security Implication of AI-altered Information Environments.,” United Nations University Centre for Policy Research, 2024. [Online]. Available: https://unu.edu/publication/disinformation-and-peacebuilding-sub-saharan-africa
[51] C. Bjola, “AI for development: implications for theory and practice,” *Oxford Development Studies*, vol. 50, no. 1, pp. 78–90, Aug. 2021, doi: 10.1080/13600818.2021.1960960.
[52] H. Ashby, “A role for AI in peacebuilding,” United States Institute of Peace. [Online]. Available: https://www.usip.org/publications/2023/12/role-ai-peacebuilding
[53] R. Muggah and I. Szabo, “Artificial intelligence will entrench global inequality,” *Foreign Policy*, May 29, 2023. [Online]. Available: https://foreignpolicy.com/2023/05/29/ai-regulation-global-south-artificial-intelligence/
[54] A. Wang and P. Triolo, “There can be no winners in a US-China AI arms race,” *MIT Technology Review*, Jan. 2025, [Online]. Available: https://www.technologyreview.com/2025/01/21/1110269/there-can-be-no-winners-in-a-us-china-ai-arms-race/
[55] P. Hidalgo-Sanchis, “UNU Macau AI Conference, AI for All: Bridging Divides, Building a Sustainable Future. Policy directions distilled by UNU Macau.” United Nations University, 2024. [Online]. Available: https://unu.edu/sites/default/files/2024-08/UNU_AI_ConferencePolicyReport_0806.pdf
[56] H. A. Ünver, “Artificial Intelligence (AI) and Human Rights: Using AI as a Weapon of Repression and Its Impact on Human Rights,” European Parliament, May 2024. [Online]. Available: https://www.europarl.europa.eu/thinktank/en/document/EXPO_IDA(2024)754450
[57] “Emerging Technologies and Terrorism: An American Perspective.” Army War College Press, Apr. 18, 2024. [Online]. Available: https://press.armywarcollege.edu/monographs/967
[58] F. Ogenga and A. Stanley, “Regulating artificial intelligence in Africa: Strategies and insights from Kenya, Ghana, and the African Union.” [Online]. Available: https://www.wilsoncenter.org/blog-post/regulating-artificial-intelligence-africa-strategies-and-insights-kenya-ghana-and-african

[59] C. Arun, “Transnational AI and corporate imperialism,” Carnegie Endowment for International Peace. [Online]. Available: https://carnegieendowment.org/research/2024/10/transnational-ai-and-corporate-imperialism?lang=en
[60] K. McQue, L. Martins, A. Bhattacharya, and C. Du Plessis, “The global struggle over how to regulate AI,” *Rest of World*, Jan. 21, 2025. [Online]. Available: https://restofworld.org/2025/global-ai-regulation-big-tech/
[61] E. Tauchnitz, “Alte und neue Haftungslücken im digitalen Zeitalter,” *Stratos (Military Science Journal of the Swiss Army)*, no. Special Issue, pp. 136–146, Nov. 2024.
[62] J. Cook, “How Israel uses an AI genocide program to obliterate Gaza,” *Washington Report on Middle East Affairs*, Dec. 15, 2023. [Online]. Available: https://www.wrmea.org/israel-palestine/how-israel-uses-an-ai-genocide-program-to-obliterate-gaza.html
[63] V. D’Evereux, “Israeli military artificial intelligence, its possible use in the war in Gaza,” *Obrana a strategie*, vol. 24, no. 1, pp. 125–142, Jun. 2024, doi: 10.3849/1802-7199.24.2024.01.125-142.
[64] Y. Abraham, “A mass assassination factory’: Inside Israel’s calculated bombing of Gaza,” *+972 Magazine*, Nov. 30, 2023. [Online]. Available: https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/
[65] T. E. Minu, “How Israel uses AI in Gaza—and what it might mean for the future of warfare,” *Time*, Dec. 18, 2024. [Online]. Available: https://time.com/7202584/gaza-ukraine-ai-warfare/
[66] Y. Abraham, “‘Lavender’: The AI machine directing Israel’s bombing spree in Gaza,” *+972 Magazine*, Apr. 03, 2024. [Online]. Available: https://www.972mag.com/lavender-ai-israeli-army-gaza/
[67] D. Kirichenko, “The rush for AI-enabled drones on Ukrainian battlefields,” Lawfare. [Online]. Available: https://www.lawfaremedia.org/article/the-rush-for-ai-enabled-drones-on-ukrainian-battlefields
[68] S. Bendett and D. Kirichenko, “Battlefield drones and the accelerating autonomous arms race in Ukraine,” Modern War Institute. [Online]. Available: https://mwi.westpoint.edu/battlefield-drones-and-the-accelerating-autonomous-arms-race-in-ukraine/
[69] L. Leffer, “AI chatbots will never stop hallucinating,” *Scientific American*, Apr. 05, 2024. [Online]. Available: https://www.scientificamerican.com/article/chatbot-hallucinations-inevitable/
[70] High-level Advisory Body on Artificial Intelligence, “Governing AI for Humanity: Final Report,” United Nations, Sep. 2024. [Online]. Available: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
[71] C. Zelizer, Ed., *Integrated Peacebuilding: Innovative Approaches to Transforming Conflict.* NY: Routledge, 2013.
[72] International Alert, “International Alert’s Approach to Peacebuilding,” Feb. 2023. [Online]. Available: https://www.international-alert.org/app/uploads/2023/02/International-Alerts-Approach-To-Peacebuilding-EN-2023.pdf
[73] B. Goldberg *et al.*, “AI and the future of digital public squares,” Dec. 13, 2024, *Arxiv preprint*. [Online]. Available: https://arxiv.org/abs/2412.09988
[74] L. Schirch, “The case for designing tech for social cohesion: The limits of content moderation and tech regulation,” *Yale Journal of Law & the Humanities.*, vol. 34, no. 3, Jan. 2023, [Online]. Available: https://openyls.law.yale.edu/handle/20.500.13051/18411
[75] Organization for Economic Cooperation and Development, “Peace and Official Development Assistance,” OECD, 37, Oct. 2023. [Online]. Available: https://www.oecd.org/en/publications/peace-and-official-development-assistance_fccfbffc-en.html
[76] Organization for Economic Cooperation and Development, *States of Fragility 2022*. OECD, 2022. doi: 10.1787/c7fedf5e-en.
[77] United Nations Peacebuilding Support Office, “UN Peacebuilding: An Orientation,” 2010. [Online]. Available: https://www.un.org/peacebuilding/sites/www.un.org.peacebuilding/files/documents/peacebuilding_orientation.pdf
[78] M. Wallace, *From Principle to Practice: A User’s Guide to Do No Harm. Revised 2016.* [Online]. Available: https://www.principletopractice.org/docs/From%20Principle%20to%20Practice%20Book.pdf
[79] Organization for Economic Cooperation and Development, *States of Fragility 2025*. in States of Fragility. OECD Publishing, 2025. doi: 10.1787/81982370-en.

[80] Stockholm International Peace Research Institute, “Global military spending surges amid war, rising tensions and insecurity,” Apr. 2024. [Online]. Available: https://www.sipri.org/media/press-release/2024/global-military-spending-surges-amid-war-rising-tensions-and-insecurity
[81] United Nations, “Fifth Committee Approves $5.59 Billion Budget for 14 Peacekeeping Operations, Service Centres, Headquarters Support Staff, Concluding Resumed Session.” Jun. 21, 2024. [Online]. Available: https://press.un.org/en/2024/gaab4463.doc.htm
[82] R. Davies, “Burden-shedding: the unravelling of the OECD aid consensus - Devpolicy Blog from the Development Policy Centre.” Mar. 07, 2025. [Online]. Available: https://devpolicy.org/burden-shedding-the-unravelling-of-the-oecd-aid-consensus-20250307/
[83] International Telecommunication Union, “United Nations Activities on Artificial Intelligence (AI) 2023.” 2024. [Online]. Available: https://www.itu.int/hub/publication/s-gen-unact-2023/
[84] UN Innovation Network, “Generative AI.” [Online]. Available: https://www.uninnovation.network/un-group-pages/generative-ai
[85] African Union, “The Genesis of ECOSOCC.” [Online]. Available: https://ecosocc.au.int/en/about/overview
[86] B. Davletov, U. Kalkar, M. Ragnet, and S. Verhulst, “PeaceTech Topic Map: A Research Base for an Emerging Field,” Global PeaceTech Hub, Jan. 2023. [Online]. Available: https://www.globalpeacetech.org/global-peacetech-hub-govlab-research-peace-tech-field/
[87] Amani Africa, “Artificial Intelligence and its impact on peace, security and governance.” [Online]. Available: https://amaniafrica-et.org/artificial-intelligence-and-its-impact-on-peace-security-and-governance/
[88] “What does the Council do?,” Council on Technology and Social Cohesion. [Online]. Available: https://techandsocialcohesion.org/what-2/
[89] UNESCO, “Recommendation on the Ethics of Artificial Intelligence,” 2022. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137
[90] United Nations, “Global Digital Compact.” Sep. 2024. [Online]. Available: https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf
[91] Build Up, “We transform conflict in the digital age.” [Online]. Available: https://howtobuildup.org/
[92] Alliance for Peacebuilding, “Digital peacebuilding community of practice.” [Online]. Available: https://www.allianceforpeacebuilding.org/digital-peacebuilding-cop
[93] I. Luna-Pla, “Network analysis in peace and state building: revealing power elites,” *Data & Policy*, vol. 6, Apr. 2024, doi: 10.1017/dap.2024.9.
[94] E. Pauwels, “Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention,” Global Center on Cooperative Security, Policy Brief, Sep. 2020. [Online]. Available: https://globalcenter.org/resource/artificial-intelligence-and-data-capture-technologies-in-violence-and-conflict-prevention/
[95] S. Spencer, “Humanitarian AI: The hype, the hope and the future,” Humanitarian Practice Network, 85, Nov. 2021. [Online]. Available: https://odihpn.org/wp-content/uploads/2021/11/HPN-Network-Paper_AI_web_181121.pdf
[96] J. Bryant, “Digital mapping and inclusion in humanitarian response,” Humanitarian Policy Group, Oct. 2021. [Online]. Available: https://media.odi.org/documents/Digital_inclusion_synthesis.pdf
[97] M. Croicu and S. P. von der Maase, “From newswire to nexus: Using text-based actor embeddings and transformer networks to forecast conflict dynamics,” Jan. 07, 2025, *arXiv*. doi: 10.48550/ARXIV.2501.03928.
[98] C. Raleigh, R. Kishi, and A. Linke, “Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices,” *Humanities and Social Sciences Communications*, vol. 10, no. 1, Feb. 2023, doi: 10.1057/s41599-023-01559-4.
[99] S. Rejali and Y. Heiniger, “The role of digital technologies in humanitarian law, policy and action: Charting a path forward,” *International Review of the Red Cross*, vol. 102, no. 913, pp. 1–22, Apr. 2020, doi: 10.1017/s1816383121000114.
[100] C. Martin-Shields, “The technologist’s dilemma: ethical challenges of using crowdsourcing technology in conflict and disaster-affected regions,” *Georgetown Journal of International Affairs*, vol. 14, no. 2, pp. 157–163, 2013.
[101] K. Peach, A. Berditchevskaia, I. Gill, O. Whittington, E. Malliaraki, and N. Hussein, “Collective Crisis Intelligence for Frontline Humanitarian Response,” Nesta, Sep. 2021. [Online]. Available: https://www.nesta.org.uk/report/collective-crisis-intelligence-frontline-humanitarian-response/

[102] A. T. Hirblinger, “When the digits don’t add up: Research strategies for post-digital peacebuilding,” *Cooperation and Conflict*, vol. 59, no. 3, pp. 425–446, Aug. 2023, doi: 10.1177/00108367231184727.
[103] United Nations Environment Programme, “Digital Technologies for Environmental Peacebuilding: Horizon Scanning of Opportunities and Risks,” 2024. [Online]. Available: https://digitallibrary.un.org/record/4058938?v=pdf
[104] H. Hegre, N. W. Metternich, H. M. Nygård, and J. Wucherpfennig, “Introduction,” *Journal of Peace Research*, vol. 54, no. 2, pp. 113–124, Feb. 2017, doi: 10.1177/0022343317691330.
[105] N. N. Taleb, *The Black Swan: The Impact of The Highly Improbable*. New York: Random House, 2016.
[106] B. Panic, Advancing AI for effective crisis early warning early action – Insights from PRIO AI days. [Online]. Available: https://www.linkedin.com/pulse/advancing-ai-effective-crisis-early-warning-action-insights-panic-qssve
[107] A. Beduschi, “Harnessing the potential of artificial intelligence for humanitarian action: Opportunities and risks,” *International Review of the Red Cross*, vol. 104, no. 919, pp. 1149–1169, Apr. 2022, doi: 10.1017/s1816383122000261.
[108] A. T. Hirblinger, J. M. Hansen, K. Hoelscher, Å. Kolås, K. Lidén, and B. O. Martins, “Digital peacebuilding: A framework for critical–reflexive engagement,” *International Studies Perspectives*, vol. 24, no. 3, pp. 265–284, Nov. 2022, doi: 10.1093/isp/ekac015.
[109] Uppsala University, “About UCDP.” [Online]. Available: https://www.uu.se/en/department/peace-and-conflict-research/research/ucdp/about-ucdp
[110] H. Hegre, H. M. Nygård, and P. Landsverk, “Can we predict armed conflict? How the first 9 years of published forecasts stand up to reality,” *International Studies Quarterly*, vol. 65, no. 3, pp. 660–668, Sep. 2021, doi: 10.1093/isq/sqaa094.
[111] H. Hegre, M. Croicu, K. Eck, and S. Hogbladh, “Introducing the UCDP Candidate Events Dataset,” *Research & Politics*, vol. 7, no. 3, Jul. 2020, doi: 10.1177/2053168020935257.
[112] H. Hegre, J. Karlsen, H. Mokleiv, H. Strand, and H. Urdal, “Forecasting Fatalities in Armed Conflict: ViEWS Forecasts for April 2022–March 2025,” Uppsala University, May 2022. [Online]. Available: https://uu.diva-portal.org/smash/get/diva2:1665945/FULLTEXT01.pdf
[113] Armed Conflict Location and Event Data, “Annual Report 2021.” [Online]. Available: https://acleddata.com/about-acled/annual-report-2021/#s6
[114] E. Albrecht, “Predictive Technologies in Conflict Prevention: Practical and Policy Considerations for the Multilateral System,” United Nations University Centre for Policy Research, Jun. 2023. [Online]. Available: https://unu.edu/sites/default/files/2023-09/predictive_technologies_conflict_prevention_.pdf
[115] K. Singh, “Rights advocates concerned by reported US plan to use AI to revoke student visas.,” *Reuters*, Mar. 06, 2025. [Online]. Available: https://www.reuters.com/technology/artificial-intelligence/us-use-ai-revoke-visas-students-perceived-hamas-supporters-axios-reports-2025-03-06/
[116] S. Khan, R. Baker, C. Howarth, I. Baker, N. Adler, and S. Verhulst, “Children on the Move: Using Satellite Data Analysis in Conflict/Famine-Affected Areas,” Harvard Humanitarian Initiative, Nov. 2019. [Online]. Available: http://pubdocs.worldbank.org/en/709811591367809253/01-Lessons-Learned-Public-Report-11-20-19.pdf
[117] A. Hunt and D. Specht, “Crowdsourced mapping in crisis zones: Collaboration, organisation and impact,” *Journal of International Humanitarian Action*, vol. 4, no. 1, Jan. 2019, [Online]. Available: https://jhumanitarianaction.springeropen.com/articles/10.1186/s41018-018-0048-1
[118] A. R. Shahid, “The Impact of Crowdmapping on Humanitarian Response: A Structurational Analysis,” Doctoral Thesis, School of Business and Management, Royal Halloway University, London, 2016. [Online]. Available: https://pure.royalholloway.ac.uk/en/publications/the-impact-of-crowdmapping-on-humanitarian-response-a-structurati
[119] C. Morgan, “HOT Impact Report 2022–2023,” Humanitarian OpenStreetMap Team, 2023. [Online]. Available: https://www.hotosm.org/impact-report-2022-2023/
[120] F. Oliva and B. Mcquinn, “Harnessing the potential of human-in-the-loop artificial intelligence for risk anticipation and violence prevention,” UNDP Global Policy Network Brief, Aug. 2023. [Online]. Available: https://www.undp.org/publications/dfs-harnessing-potential-human-loop-artificial-intelligence-risk-anticipation-and-violence-prevention

[121] A. Sarfati, “New Technologies and the Protection of Civilians in UN Peace Operations,” International Peace Institute, Sep. 2023. [Online]. Available: https://www.jstor.org/stable/resrep53082
[122] Planet Labs PBC, “Ukraine photo story.” [Online]. Available: https://www.planet.com/ukraine-photo-story/
[123] R. Avtar *et al.*, “Remote sensing for international peace and security: Its role and implications,” *Remote Sensing*, vol. 13, no. 3, p. 439, Jan. 2021, doi: 10.3390/rs13030439.
[124] K. Jauer, “What is DISHA? - Disha.” [Online]. Available: https://disha.unglobalpulse.org/about/
[125] L. Bromley, K. Jauer, and Y. Matias, “AI from Google Research and UN boosts humanitarian disaster response: Wider coverage, faster damage assessments,” disha - UN Global Pulse. [Online]. Available: https://disha.unglobalpulse.org/ai-from-google-research-and-un-boosts-humanitarian-disaster-response-wider-coverage-faster-damage-assessments/
[126] M. T. Mandefro, “Displacement forecast odels (Danish Refugee Council),” United Kingdom Humanitarian Innovation Hub. [Online]. Available: https://www.ukhih.org/news/displacement-forecast-models-danish-refugee-council/
[127] Data Entry and Exploration Platform, “DEEP.” [Online]. Available: https://www.thedeep.io/
[128] Urban Foresight, “Deep Dive: A comprehensive review of DEEP – the Data Entry and Exploration Platform.,” Nov. 2022. [Online]. Available: https://urbanforesight.com/wp-content/uploads/2022/11/Comprehensive-review-of-the-DEEP_Urban-Foresight_11-Nov-2022.pdf
[129] E. Albrecht, E. Fournier-Tombs, and R. Brubaker, “Disinformation and Peacebuilding in Sub-Saharan Africa: Security Implication of AI-altered Information Environments,” United Nations University - Center for Policy Research, 2024. [Online]. Available: https://unu.edu/publication/disinformation-and-peacebuilding-sub-saharan-africa
[130] M. H. Tessler *et al.*, “AI can help humans find common ground in democratic deliberation,” *Science*, vol. 386, no. 6719, Oct. 2024, doi: 10.1126/science.adq2852.
[131] A. T. Hirblinger, “When mediators need machines (and vice versa): Towards a research agenda on hybrid peacemaking intelligence,” *International Negotiation*, vol. 28, no. 1, pp. 94–125, Jan. 2022, doi: 10.1163/15718069-bja10050.
[132] H. Alessa, “The role of Artificial Intelligence in Online Dispute Resolution: A brief and critical overview,” *Information & Communications Technology Law*, vol. 31, no. 3, pp. 319–342, Jun. 2022, doi: 10.1080/13600834.2022.2088060.
[133] D. Masood Alavi, M. Wahlisch, C. Irwin, and A. Konya, “Using artificial intelligence for peacebuilding,” *Journal of Peacebuilding and Development*, vol. 17, no. 2, pp. 239–243, May 2022, doi: 10.1177/15423166221102757.
[134] A. Konya, L. Schirch, C. Irwin, and A. Ovadya, “Democratic policy development using collective dialogues and AI,” Nov. 03, 2023, *arXiv*. doi: 10.48550/ARXIV.2311.02242.
[135] C. Irwin, D. Masood Alavi, M. Waehlisch, and A. Konya, “Using Artificial Intelligence in Peacemaking: The Libya Experience,” presented at the WAPOR 74th Annual Conference, Virtual Conference, Online, Nov. 02, 2021. [Online]. Available: https://peacepolls.etinu.net/peacepolls/documents/009260.pdf
[136] L. Schirch, “The Digital Space and Peace Processes: A Thought Piece,” Principles for Peace, May 2022. [Online]. Available: https://ict4peace.org/wp-content/uploads/2022/06/P4P-Digital-Space-and-Peace-Processes-v2.pdf
[137] A. Konya *et al.*, “Using collective dialogues and AI to find common ground between Israeli and Palestinian peacebuilders,” Mar. 07, 2025, *arXiv*: arXiv:2503.01769. doi: 10.48550/arXiv.2503.01769.
[138] Alliance for Middle East Peace, “AI Pulse.” [Online]. Available: https://www.allmep.org/ai-pulse/
[139] S. Thompson and A. Piirtola, “Artificial Intelligence and Peace - The Case of Digital Dialogues in Sudan,” CMI, Feb. 2024. [Online]. Available: https://cmi.fi/wp-content/uploads/2024/02/CMI_INSIGHT_2024_sudan.pdf
[140] B. Marnette and C. McKenzie, “Talk to the City: an open-source AI tool for scaling deliberation.” [Online]. Available: https://ai.objectives.institute/blog/talk-to-the-city-an-open-source-ai-tool-to-scale-deliberation
[141] F. Ogenga, “Social Media Literacy, Ethnicity and Peacebuilding in Kenya,” Toda Peace Institute, Policy Brief 60, Nov. 2019. [Online]. Available: https://toda.org/assets/files/resources/policy-briefs/t-pb-60_fredrick-ogenga_social-media-literacy.pdf
[142] H. Puig Larrauri, “Societal Divides as a Taxable Negative Externality of Digital Platforms: An exploration of the rationale for regulating algorithmically mediated platforms differently,” Ashoka & Build Up, Mar. 2023. [Online].

Available: https://www.next-now.org/sites/default/files/2023-03/Societal%20Divides%20as%20a%20taxable%20negative%20externality%20of%20digital%20platforms_0.pdf
[143] A. Ovadya and L. Thorburn, “Bridging Systems: Open problems for countering destructive divisiveness across ranking, recommenders, and governance,” Knight First Amendment Institute at Columbia University, Oct. 2023. [Online]. Available: https://knightcolumbia.org/content/bridging-systems
[144] A. Ovadya, “Bridging-Based Ranking How Platform Recommendation Systems Might Reduce Division and Strengthen Democracy,” Belfer Center for Science and International Affairs, Harvard Kennedy School., May 2022. [Online]. Available: https://www.belfercenter.org/publication/bridging-based-ranking.
[145] T. Bernard, “Can tech promote social cohesion?,” *Tech Policy Press*, Apr. 03, 2023. [Online]. Available: https://techpolicy.press/can-tech-promote-social-cohesion
[146] T. Bernard, “Social cohesion technologies and online projects,” *Tech Policy Press*, Apr. 2023, [Online]. Available: https://www.techpolicy.press/social-cohesion-technologies-and-online-projects/
[147] Pulse Lab Kampala, United Nations Global Pulse, “Using Machine Learning to Analyse Radio Content in Uganda: Opportunities for Sustainable Development and Humanitarian Action.” Pulse Lab Kampala, United Nations Global Pulse., Sep. 2017. [Online]. Available: https://unsdg.un.org/sites/default/files/Using-machine-learning-radio-content-uganda.pdf
[148] F. Ogenga, “Exploiting pan-African digital peacebuilding infrastructures in Kenya: The case of Maskani.” [Online]. Available: https://kujenga-amani.ssrc.org/2024/12/17/exploiting-pan-african-digital-peacebuilding-infrastructures-in-kenya-the-case-of-maskani-2/
[149] L. Schirch, “Digital information, conflict and democracy,” in *Social Media Impacts on Conflict and Democracy*, 1st ed., L. Schirch, Ed., New York: Routledge, 2021, pp. 21–42. doi: 10.4324/9781003087649-2.
[150] S. John, “Digital peacebuilding in Africa: trends, challenges, opportunities,” *African Journal of Peace and Conflict Studies*, vol. 13, no. 3, Dec. 2024, [Online]. Available: https://journals.co.za/doi/abs/10.31920/2634-3665/2024/v13n3a4
[151] Jigsaw, “Exploring how AI might enhance our digital public squares.” [Online]. Available: https://medium.com/jigsaw/exploring-how-ai-might-enhance-our-digital-public-squares-ec60eb702887
[152] N. Davis, “AI mediation tool may help reduce culture war rifts, say researchers,” *The Guardian*, Oct. 17, 2024. [Online]. Available: https://www.theguardian.com/technology/2024/oct/17/ai-mediation-tool-may-help-reduce-culture-war-rifts-say-researchers
[153] MIT AI Risk Repository, “Risk Classification.” [Online]. Available: https://airisk.mit.edu/ai-incident-tracker/risk-classification
[154] N. Jude and A. Matamoros-Fenandez, “Community notes and its narrow understanding of disinformation,” Tech Policy Press. [Online]. Available: https://www.techpolicy.press/community-notes-and-its-narrow-understanding-of-disinformation/
[155] A. Mahdawi, “AI-generated ‘slop’ is slowly killing the internet, so why is nobody trying to stop it?,” *The Guardian*. [Online]. Available: https://www.theguardian.com/global/commentisfree/2025/jan/08/ai-generated-slop-slowly-killing-internet-nobody-trying-to-stop-it
[156] “Mercado Libre introduces Verdi, an AI developer platform powered by GPT-4o,” OpenAI. [Online]. Available: https://openai.com/index/mercado-libre/
[157] C. Rule, “Online Dispute Resolution Moves From E-Commerce to the Courts,” the Pew Charitable Trusts. [Online]. Available: https://www.pewtrusts.org/en/research-and-analysis/articles/2019/06/04/online-dispute-resolution-moves-from-e-commerce-to-the-courts
[158] M. H. Tessler *et al.*, “AI can help humans find common ground in democratic deliberation,” *Science*, vol. 386, no. 6719, Oct. 2024, doi: 10.1126/science.adq2852.
[159] K. Manson, “The US military is taking generative AI out for a spin,” *Bloomberg*, Jul. 05, 2023. [Online]. Available: https://www.bloomberg.com/news/newsletters/2023-07-05/the-us-military-is-taking-generative-ai-out-for-a-spin?embedded-checkout=true%5D
[160] J.-P. Rivera, G. Mukobi, A. Reuel, M. Lamparth, C. Smith, and J. Schneider, “Escalation risks from language models in military and diplomatic decision-making,” Jan. 07, 2024, *arXiv*. doi: 10.48550/ARXIV.2401.03408.

[161] Omdena, “Leveraging LLMs to enhance multilateral negotiations with AI-powered decision-making.” [Online]. Available: https://www.omdena.com/projects/leveraging-llms-in-building-a-global-negotiations-decision-making-framework
[162] T. Byrnes, “Harnessing the Transformative Potential of Generative AI for Humanitarian Multi-Purpose Cash Assistance Opportunities, Risks, Barriers, and Recommendations,” Market Impact, May 2024. [Online]. Available: https://www.calpnetwork.org/es/publication/harnessing-the-transformative-potential-of-generative-ai-for-humanitarian-multi-purpose-cash-assistance-opportunities-risks-barriers-and-recommendations/
[163] G. Connors and E. Baumhofer, “Peacebuilding And Disinformation: Taking Stock and Planning Ahead.” Berghof Foundation & Platform Peaceful Conflict Transformation, 2025. [Online]. Available: https://berghof-foundation.org/library/peacebuilding-and-disinformation
[164] Stimson Center, “Future of International Cooperation Report 2023: Building Shared Futures: Innovating Governance for Global and Regional Problem-Solving,” Sep. 2023. [Online]. Available: https://www.stimson.org/2023/future-international-cooperation-report-2023/
[165] Inter-American Development Bank, “Publications.” [Online]. Available: https://publications.iadb.org/en
[166] Arist, “About Arist.” [Online]. Available: https://www.arist.co/about
[167] Unschooler, “AI-training for your company business metrics.” [Online]. Available: https://unschooler.me/
[168] K. Butterby and N. Lombard, “Developing a chatbot to support victim-survivors who are subjected to domestic abuse: considerations and ethical dilemmas,” *Journal of Gender-Based Violence*, vol. 9, no. 1, pp. 153–161, Jan. 2025, doi: 10.1332/23986808Y2024D000000038.
[169] H. Hussain, “Why Chayn took down its chatbot in 2020 and what we’ve learned about building culturally-aware chatbots since,” Medium. [Online]. Available: https://blog.chayn.co/why-chayn-took-down-its-chatbot-in-2020-and-what-weve-learned-about-culturally-aware-chatbots-a9587cf80df8
[170] M. V. Heinz *et al.*, “Randomized Trial of a Generative AI Chatbot for Mental Health Treatment,” *NEJM AI*, vol. 2, no. 4, Mar. 2025, doi: 10.1056/AIoa2400802.
[171] V. Pietromarchi, “Can AI mediate conflict better than humans?,” *Al Jazeera*, Feb. 29, 2024. [Online]. Available: https://aje.io/jww93m
[172] C. Wilmot, “Can AI bring peace to the Middle East?,” *The Bureau of Investigative Journalism*, Dec. 19, 2024. [Online]. Available: https://www.thebureauinvestigates.com/stories/2024-12-19/can-ai-bring-peace-to-the-middle-east

ANNEX

Table 2. Summary of Data Sources for the Technical Report

Category	Details
Primary Search Areas	AI & Peacebuilding Machine Learning for Conflict Prevention/Resolution AI in Early Warning Systems Peace Technology Digital Peacebuilding & Mediation Predictive Analysis for Conflict
Key AI applications	Conflict Prediction & Monitoring Natural Language Processing Sentiment & Social Media Analysis Violence Prevention Digital Peace Infrastructure
Implementation Contexts	Ethical and Responsible AI Local & Community Based Peacebuilding Post-Conflict Recovery Policy & Governance
Key Data Sources	Academic Research International Organizations Think Tanks and Networks Civil Society and NGOs Private Sector Foundations and Advocacy Groups Media Outlets

Source: International Panel on the Information Environment

ACKNOWLEDGMENTS

Contributors

Drafting authors: Craig Zelizer (Consulting Scientist, Colombia), Fredrick Ogenga (Panel Chair, Kenya), Lisa Schirch (Panel Vice Chair, USA) and Evelyne Tauchnitz (Panel Vice Chair, Switzerland), Sebastián Valenzuela (IPIE Chief Science Officer and Chair of the Science & Methodology Committee, Chile), Philip Howard (IPIE President and CEO, Canada/UK). Independent General Reviews: Joseph Aylett-Bullock, Michele Giovanardi and Branka Panic. Fact-checking: Heidi Schultz. Design: Domenico Di Donna. Copyediting: Beverley Sykes. We gratefully acknowledge support from the IPIE Secretariat: Lola Gimferrer, Egerton Neto, Wiktoria Schulz, Donna Seymour, Anna Staender, and Alex Young.

Funders

The International Panel on the Information Environment (IPIE) gratefully acknowledges the support of its funders. For a full list of funding partners please visit www.IPIE.info. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the IPIE and do not necessarily reflect the views of the funders.

Declaration of Interests

IPIE reports are developed and reviewed by a global network of research affiliates and consulting scientists who constitute focused Scientific Panels and contributor teams. All contributors and reviewers complete declarations of interests, which are reviewed by the IPIE at the appropriate stages of work.

Preferred Citation

An IPIE *Summary for Policymakers* provides a high-level precis of the state of knowledge and is written for a broad audience. An IPIE *Synthesis Report* makes use of scientific meta-analysis techniques, systematic review, and other tools for evidence aggregation, knowledge generalization, and scientific consensus building, and is written for an expert audience. An IPIE *Technical Paper* addresses particular questions of methodology, or provides a policy analysis on a focused regulatory problem. All reports are available on the IPIE website (www.IPIE.info).

This document should be cited as:

International Panel on the Information Environment [C. Zelizer, F. Ogenga, L. Schirch, E. Tauchnitz, P. N. Howard, S. Valenzuela (eds.)], “Artificial Intelligence and Peacebuilding: Opportunities and Challenges,” Zurich, Switzerland: IPIE, 2025. Technical Paper, TP2025.3, doi: 10.61452/RNGW7145.

Copyright Information

ABOUT THE IPIE

The International Panel on the Information Environment (IPIE) is an independent and global science organization providing scientific knowledge about the health of the world's information environment. Based in Switzerland, the IPIE offers policymakers, industry, and civil society actionable scientific assessments about threats to the information environment, including AI bias, algorithmic manipulation, and disinformation. The IPIE is the only scientific body systematically organizing, evaluating, and elevating research with the broad aim of improving the global information environment. Hundreds of researchers worldwide contribute to the IPIE's reports.

For more information, please contact the International Panel on the Information Environment (IPIE), secretariat@IPIE.info. Seefeldstrasse 123, P.O. Box, 8034 Zurich, Switzerland.

International Panel on
the Information
Environment

Seefeldstrasse 123
P.O. Box 8034 Zurich
Switzerland