



**Alliance for
Peacebuilding**

PARADIGM SHIFT

Would you say that you dream about the future?



Designing AI

for Conflict Prevention & Peacebuilding

Alliance for Peacebuilding | October 2023



United Nations Security Council meeting on Artificial Intelligence in New York City, May 2023. [Rory Arnold/No 10 Downing Street]

Designing AI for Conflict Prevention & Peacebuilding

Nick Zuroski, Manager - Policy & Advocacy, AfP

Megan Corrado, Director - Policy & Advocacy, AfP

Liz Hume, Executive Director, AfP

The growing prominence of artificial intelligence (AI) poses, as President Biden stated, “[enormous potential and enormous danger](#)” for global peace and security. The peacebuilding field, technology engineers, and policymakers must work together to minimize the dangers of AI and maximize AI’s potential to prevent and reduce violent conflict, violence, and fragility and build sustainable peace globally. AI technology can vastly [increase humanity’s capacity to make discoveries through machine learning](#) and coherently [summarize numerous text documents](#). However, AI can also [easily generate](#) mis and disinformation at a

massive scale and, as a worst-case scenario, [support deadly automated weaponry](#) beyond human control. The peacebuilding field needs to raise awareness and educate how AI technology can positively and negatively impact global peace and security, and governments and multilateral organizations must urgently adopt robust AI standards and norms that prevent and reduce violent conflict, violence, and fragility and build sustainable peace. Additionally, peacebuilders should work with the technology sector to develop AI-powered tools that “translate” divisive rhetoric into peaceful language and thus promote social cohesion and peacebuilding.

What is AI?

While no one definition of AI exists, the Organization for Economic Cooperation and Development (OECD) [defines](#) it as “the ability of machines and systems to acquire and apply knowledge, and to carry out intelligent behavior.”¹ There are various types of AI, each with unique applications, including [speech recognition](#), [image/text analysis and generation](#), [personalized education experiences](#), [autonomous vehicles](#), [improved accuracy of medical diagnosis](#), [data summarization](#), and [chatbots](#). [Machine learning](#) uses AI algorithms to recognize patterns in data already categorized by humans and then make informed predictions based on new, emerging data—for instance, [using someone’s evolving purchase and browsing history](#) to recommend personalized product advertisements online. [Deep learning](#) is a more advanced process similar to machine learning in its goals. It uses complex artificial neural networks modeled after the human brain to analyze uncategorized data, requiring less human intervention, such as how self-driving cars can analyze data to detect stop signs. Generative AI creates content across different formats, such as audio, videos, images, and text.

Despite decades of advances in AI development, an inflection point came in November 2022, when OpenAI [released GPT 3.5](#), followed by the [improved GPT-4](#) in March 2023. ChatGPT (generative pre-trained transformer) is a form of AI known as a [large language model](#) trained on vast amounts of online data to predictively generate text based on a text prompt. ChatGPT can be used, for example, to write an email in the style of a particular author, summarize a long document in a few paragraphs, or [even code](#). ChatGPT was a breakthrough for public recognition of AI because of its [vast application to many sectors](#), [lack of mandatory subscription costs](#), and [easy-to-use public interface](#). In late September 2023, OpenAI announced it would soon release an update for ChatGPT to [enable voice conversations and image analysis](#).

1. IBM categorizes AI as [weak or strong](#), with weak AI designed for specific tasks and strong AI (still theoretical and otherwise known as artificial general intelligence or AGI) being a self-aware consciousness equal to humans capable of solving problems and learning across different scenarios. However, some researchers claim AI is [beginning to show signs of AGI](#).

There are various types of AI, each with unique applications:



Speech recognition



Image/text analysis & generation



Personalized education experiences



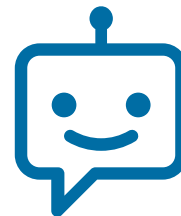
Autonomous vehicles



Data summarization



Improved accuracy of medical diagnosis



Chatbots

How AI Could Contribute to Peace & Security

While AI could spell many dangers for global peace and security, it also holds exciting potential for innovation in conflict prevention and peacebuilding. Global violent conflict and fragility are at the [highest levels in decades](#). [Twenty countries](#) are experiencing or at risk of experiencing atrocities, from [Ukraine](#) to [Myanmar](#) to [Sudan](#). In 2022, violence had a devastating economic impact, costing [\\$17.5 trillion](#), or 12.9% of global gross domestic product. To meet these unprecedented challenges, peacebuilders will need to harness AI as an amplifier of peacebuilding interventions.

AI-powered dialogue tools such as the [Remesh platform](#) can facilitate real-time conversation between large populations, which can be integrated into consultative elements of democracy and peacebuilding initiatives. The UN Department of Political and Peacebuilding Affairs (UNDPPA) uses the Remesh platform in Yemen, Libya, and Iraq to hold [AI-assisted dialogue](#) with groups of up to 1,000 citizens as part of official peace processes. AI-assisted mass dialogue can help peacebuilders facilitate real-time mediated conversations between large groups of people to build social cohesion, one day having conversations among hundreds of thousands of people in real time.

Through automation, AI can also increase the amount and real-time accuracy of data analysis in conflict-affected and fragile states to [inform early warning and response mechanisms](#) (EWER). [EWER](#) collects and analyzes data to map out conflict trends and dynamics, allowing preventative intervention by key stakeholders when warning signs of violence emerge. EWER is a [crucial tool to prevent atrocities](#), as crimes against

humanity, war crimes, genocide, and other grave human rights abuses are typically [long processes that involve early warning signs](#), such as the proliferation of dehumanizing language and violations of civic rights. By further developing AI, there are significant innovative opportunities to increase AI's capacity to [monitor massive amounts of data to conduct conflict analysis](#). AI can also automate and thus increase the accuracy and scale of EWER systems as they monitor conflict dynamics and inform preventative interventions. These bolstered analyses and systems can then better promote peacebuilding collaboration by coordinating large-scale data and conflict monitoring and increasing marginalized groups' meaningful inclusion and leadership in peace and security processes.

AI also has the potential to [refine gender analysis](#) in conflict-affected and fragile states to more clearly identify drivers of conflict and design more effective conflict prevention programming that integrates gender considerations and addresses gender-based violence (GBV). UN Women, for instance, used AI to enhance its analysis of online conflict dynamics—by looking at [gendered violent social media narratives in Bangladesh, Indonesia, Malaysia, and the Philippines](#)—to better inform peaceful counter-messaging and gender-sensitive peacebuilding interventions. AI tools used for real-time dialogue like Remesh can also help address physical and logistical obstacles to the meaningful inclusion of women in peace processes—which [increases the success and duration of peace agreements](#). AI could also [enhance GBV prevention](#) by helping governments, institutions, and local civil society organizations better collect and collaboratively analyze data on GBV incidents, which has traditionally been severely slow and uncoordinated.

AI-powered technologies could also contribute toward the Sustainable Development Goals (SDGs) in the UN 2030 Agenda for Sustainable Development. A [study](#) found that AI could enhance 134 targets within the 17 SDGs; for example, AI could improve the efficiency of climate-friendly tech, enhance economic productivity,

“What you’re doing has enormous potential and enormous danger.”

- President Biden to Silicon Valley chief executives, May 2023

and refine conflict analysis.

Importantly, AI can increase the efficiency of [underfunded](#) peacebuilding organizations. The technology could [automate operational processes such as donor tracking and data analysis in understaffed peacebuilding organizations](#). AI-generated content could also [increase the efficiency of peacebuilding communications teams](#), creating more curated campaigns across email and social media that can build and advance the peacebuilding field. This increased capacity can support more conflict prevention and peacebuilding programs—which are needed now more than ever.

The Risks of AI

However, AI also poses severe risks to peace and security. Generative AI can significantly [increase the capacity of bad actors to create disinformation](#), defined as intentionally inaccurate data used to [“deceive and... spread to do serious harm,”](#) on a large scale through text, images, videos, and audio. In March 2022, Russian hackers posted an AI-generated [deepfake](#) video of Ukrainian President Volodymyr Zelensky [supposedly urging surrender](#) in response to

Russia’s invasion to a Ukrainian news website. Though the video was quickly debunked, AI-generated disinformation has the potential to [erode trust in institutions](#) during times of crisis, which can [drive violence](#). AI also often [hallucinates](#) or inexplicably makes up bogus information and sources, which could further spread false and inaccurate information. Together, these risks could seriously undermine society’s [information resilience](#), or ability to understand information ecosystems and recognize and resist intentionally manipulative content.

Disinformation disproportionately impacts [women, marginalized groups](#), and [human rights defenders](#), as online violent language can lead to [physical attacks against these groups](#), including through [technology-facilitated gender-based violence](#) (TFGBV). Mass disinformation campaigns generated by AI can [mislead populations, foment civil unrest, and give rise to identity-based violence](#). AI can also help bad actors [undermine democracy, support rising authoritarianism, suppress human rights, close civic space, and perpetrate atrocities](#). Freedom House’s [Freedom on the Net 2023](#) found that generative AI has been used in 16 countries to spread disinformation, and in 21 countries to enforce online censorship. AI-generated content thus has the



Participant captured during the session, “Human vs Artificial Intelligence” in the congress center at the Annual Meeting 2015 of the World Economic Forum in Davos, January 2015. [Benedikt von Loebel/World Economic Forum]

potential to vastly amplify digital hate and online calls to violence, [increasing the risk of physical violence and mass atrocities](#).

[Racial and other discriminatory biases in AI technologies](#) also continue to emerge. AI engineers train AI models through massive amounts of data, including articles and blog posts scraped from the internet, many of which [contain historical bias against marginalized groups](#). Microsoft's AI chatbot Tay, designed in 2016 to have conversations with Twitter users, began to [advocate for genocide within 24 hours of its launch](#), while Google's AI-powered Google Photos [categorized pictures of Black people](#) as "gorillas." These biased data systems have the potential to further inequalities

through the decisions they make around [determining creditworthiness](#), [reviewing job applications](#), and carrying out criminal justice procedures like [bail risk assessment](#). Therefore, they can [undermine marginalized groups' security](#) in very tangible ways and [drive violence and violent conflict](#) around the world. Facial recognition AI that analyzes vast numbers of photos on the internet gives rise to [privacy concerns](#) due to its collection of images without explicit consent. Facial scanning technology can allow governments to [capture personal information of and surveil their citizens](#). For instance, the Chinese government used facial recognition AI technology [to distinguish Uyghur faces in crowds](#) to then surveil the ethnic group in Xinjiang, which resulted in mass internment and [genocide](#).



The Machine Learning Studio of the Ars Electronica Center allows to use computer vision and machine learning applications to discover how machines learn and perceive their environment. Self-driving model cars can be trained, robots with facial recognition programmed and the basic concept of a assembly line observed. [Robert Bauernhansl/Ars Electronica]

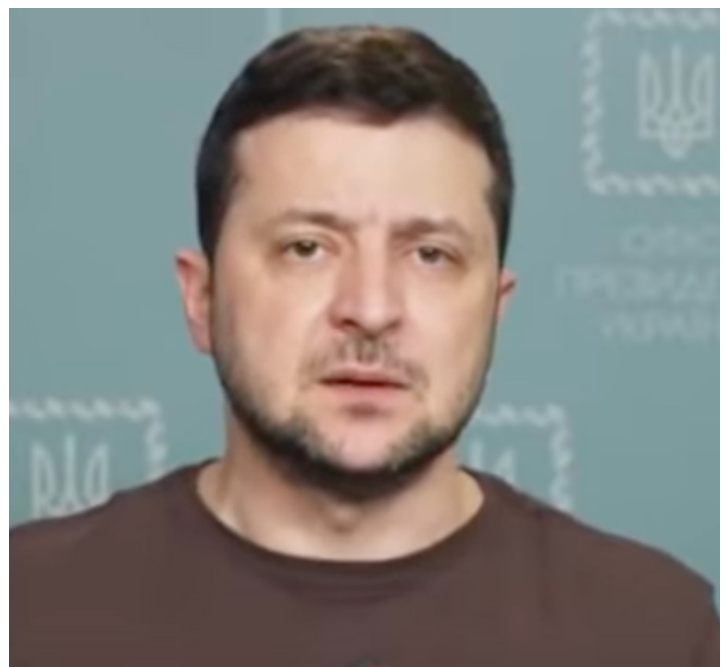
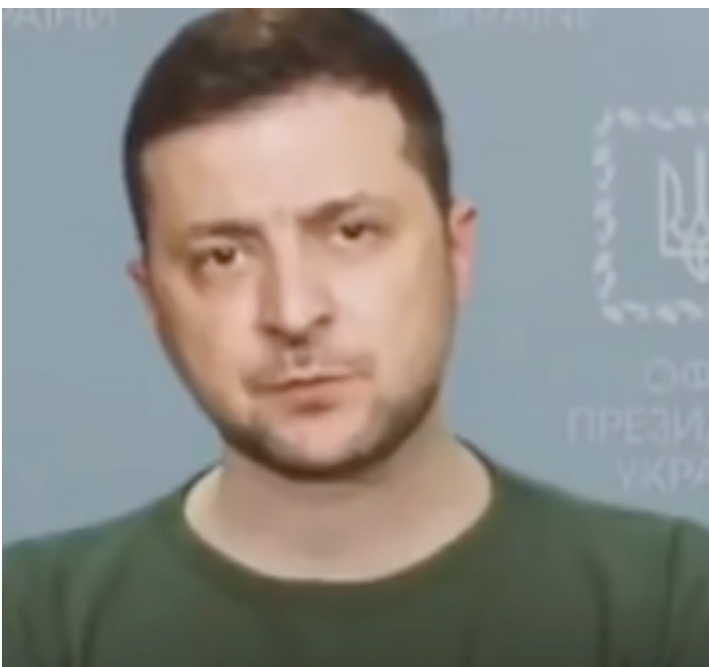
In addition, Goldman Sachs estimates that AI could automate away [300 million jobs](#)—a massive economic transformation that would profoundly impact people’s ability to earn livelihoods. Financial shocks [can drive conflict](#) as they weaken institutions, undermine the economic security of individuals, and contribute to grievances.

Experts outside of the peacebuilding community have voiced concerns about the destabilizing impacts of AI. In March 2023, an [open letter](#) signed by 1,000 tech leaders and researchers called for a six-month moratorium on developing new AI technology because of “profound risks to society and humanity,” such as propaganda, automation of all jobs, and developing “nonhuman minds that might eventually outnumber, outsmart, obsolete, and replace us.” In May 2023, the Center for AI Safety released [a one-sentence statement](#)—signed by 350 AI executives, researchers, and engineers—that urged prioritizing the risk of extinction from AI “alongside other societal-scale risks, such as pandemics and nuclear war.” This spring, Dr. Geoffrey Hinton, often referred to as “the Godfather of AI,” [quit his job](#) at Google and publicly expressed concerns about AI’s risks, including disinformation,

upending job markets, and AI becoming more intelligent than people.

Policymakers now recognize that this technology can positively and negatively disrupt life as we know it. In a [watershed Congressional hearing](#) in May 2023, Senator Richard Blumenthal noted how malign actors could use AI for destabilizing purposes by, for instance, using voice-cloning software to impersonate a U.S. Senator endorsing Ukraine’s surrender or Russian atrocities. In September 2023, Congress [began a series of closed-door listening sessions](#) with tech executives and civil society leaders to inform AI regulatory efforts later. Senator Chuck Schumer later reported to the public that every attendee of the first listening session agreed that “government is needed to play a role in AI.”

In March 2022, Russian hackers posted an AI-generated deepfake video of Ukrainian President Volodymyr Zelensky supposedly urging surrender in response to Russia’s invasion to a Ukrainian news website. The bottom left image is a deepfake and the bottom right is Zelensky’s real photo.



Recommendations

1. Governments and multilateral organizations must urgently adopt robust AI standards and norms that prevent and reduce violent conflict, violence, and fragility and build sustainable peace.

Policymakers are grappling with AI innovation and attempting to adopt requisite norms, standards, policies, and regulations. However, governments and multilateral organizations must work more urgently and robustly in partnership with peacebuilding organizations to adopt standards and norms that prevent and reduce AI's harmful impacts and ensure AI promotes conflict prevention and peacebuilding globally.

The Biden Administration has shown important recognition of AI's value and risks in policy frameworks. Still, a coordinated policy framework integrating peacebuilding and conflict prevention is vital to address AI's potential impacts on global peace and security. In October 2022, the Biden Administration released a [Blueprint for an AI Bill of Rights](#), followed by the [AI Risk Management Framework](#) in January 2023. These documents offer guidelines for organizations developing or deploying AI tools to ensure they are responsible and trustworthy and protect the rights of the American public. Their guidelines call for risk assessments to prevent harm to people or organizations, confidentiality and protection mechanisms for sensitive data, and transparency around the intended use, training data, and design of AI tools.

The U.S. has also begun institutionalizing AI in its foreign policy frameworks. In May 2022, USAID released its [Artificial Intelligence Action Plan](#), which built on USAID's 2018 report [Reflecting the Past, Shaping the Future: Making AI Work for International Development](#) and the [2020-2024 Digital Strategy](#). The action plan outlines the agency's goals to (1) commit to responsible AI in USAID programming, (2) strengthen digital ecosystems to support the responsible use of AI, and (3)

partner to shape a responsible global AI agenda. Unfortunately, the action plan fails to mention peacebuilding, only referring to "disaster prevention or response." USAID must prioritize conflict and atrocity prevention across the integration of AI within foreign assistance. The State Department [has begun to consider AI in the work of several of its bureaus and offices](#), such as countering disinformation within the Global Engagement Center and ensuring ethical and responsible use of AI in military operations in the Bureau of Arms Control, Verification, and Compliance. However, similar to those within USAID, these efforts to standardize AI in U.S. diplomacy do not meaningfully touch on conflict and atrocity prevention. The U.S. still needs to integrate

AI into its foreign assistance—primarily through a prevention lens—to develop a more innovative foreign policy that grapples with emerging technologies and promotes economic growth, good governance, and peacebuilding.

After President Biden met with tech executives in May 2023 to discuss the risks of AI, they subsequently announced [\\$140 million in investments](#) through the National Science Foundation in AI research and development, including for "transformative AI advances that are ethical, trustworthy, responsible, and serve the public good" in sectors like climate, agriculture, education,

and cybersecurity. In June 2023, Biden [met with civil society organizations](#), including the Center for Humane Technology, to discuss AI's potential risks. In July 2023, the Biden Administration [announced voluntary commitments](#) from seven leading technology companies in the U.S. to ensure AI's safe and trustworthy use. The commitments included rigorous safety and capability evaluations of AI technologies; information-sharing

Governments and multilateral organizations must work more urgently and robustly in partnership with peacebuilding organizations to adopt standards and norms that prevent and reduce AI's harmful impacts and ensure AI promotes conflict prevention and peacebuilding globally.

of trust and safety risks; cybersecurity investments; development of mechanisms that mark AI-generated content; research on societal risks posed by AI; and creating innovative AI systems for society's most significant challenges, such as climate change mitigation and adaptation. The peacebuilding community and civil society should monitor the situation to ensure crucial follow-through of these commitments in the coming months and years.

Congressional efforts have also recognized the significant impacts of AI on society, but not explicitly around foreign assistance and international conflict prevention. In June 2023, Senators Blumenthal and Hawley [proposed legislation](#) that would carve out exceptions to [Section 230](#)—which traditionally shields social media companies from liability as the publisher or speaker of information users provide on their sites—so that these companies could be sued for the spread of harmful content created by AI. Section 230 [does not protect creators or developers of content](#), and thus, generative AI companies like ChatGPT are likely to face increased

legal scrutiny as AI-produced harmful content spreads online.

Senator Michael Bennet [introduced legislation](#) that would “require relevant federal agencies to designate a senior official able to advise on the responsible use of emerging technologies like artificial intelligence” in May 2023. That same month, Senator Bennett also [introduced legislation](#) proposing a new Federal Digital Platform Commission that would regulate “content primarily generated by algorithmic processes” (such as AI technologies) and require algorithmic and public risk assessments of digital platforms of “systemic importance.” Additionally, Representative Ted Lieu has [called for a nonpartisan commission](#) outlining how Congress can regulate AI. In July 2023, Senators Elizabeth Warren and Lindsey Graham [proposed](#) creating a new U.S. regulatory commission over the technology sector, including social media and AI companies. Compared to Bennet’s legislation, Warren and Graham’s regulatory efforts would similarly focus on AI as an emerging risk area and reducing the spread of harmful online

President Joe Biden hosts a meeting on Artificial Intelligence, June 2023. [Adam Schultz/Official White House Photo]



content, but differentiate in that they have a stronger focus on anti-competitive practices and would require the largest companies under their purview to receive operating licenses.

In June 2023, Senator Chuck Schumer [announced a framework](#) for eventually creating regulation for AI, including through a series of listening sessions in the fall that began in September 2023 and engage tech companies, civil society, and academics. Members of Congress also voiced support for AI regulation in September 2023 hearings held by the Senate [Judiciary](#) and [Commerce](#) committees.

Europe is taking on a much heavier regulatory approach to AI. The European Union's (EU) [General Data Protection Regulation](#) (GDPR) prohibits AI technologies that are the sole makers of decisions that produce legal effects in someone's life or significantly affect someone similarly. The law also requires that AI technologies using the personal information of individuals obtain those people's knowledge and consent to do so, as well as provide meaningful information about the purposes of processing personal data. The EU is currently moving forward with the [AI Act](#), which would establish a classifying system that determines the level of risk an AI

technology would pose for people's health, safety, and rights and then limit high-risk uses of AI. In June 2023, the European Parliament [passed a draft version of the law](#), and a final version is expected to be passed later this fall. While the EU's AI Act and GDPR have an important focus on the need for privacy and protecting the human rights of marginalized and vulnerable groups, they lack a proper emphasis on peace, security, conflict, and atrocity prevention.

Multilateral institutions have been critical in convening discussions recognizing AI's tremendous potential and risks. Still, they must offer official commitment within international peace and security frameworks. The UN has initiated high-level discussions on AI through the September 2022 [Principles for the Ethical Use of Artificial Intelligence in the United Nations System](#). In July 2023, the UN convened the [AI for Good Global Summit](#), which aimed to ensure that the international community mitigates the potential harms of AI while also allowing AI to reach its full potential for the benefit of society. That same month, the UN Security Council [held its first session](#) on AI's threat to international peace and security. During the session, UN Secretary-General António Guterres [called for the creation of a UN watchdog agency](#) that oversees AI regulation,

European Week of Regions and Cities: "Regional artificial intelligence. The role of regions in AI development," October 2023. [EU]



AI for Good Global Summit, July 2023. [Rowan Farrell/ITU]

stating that he will convene a high-level meeting on AI and come back with options for global governance on the issue [by the end of the year](#) (with negotiations on a legally binding instrument concluded by 2026). At the 2023 United Nations General Assembly, [more than 20 countries brought up AI](#) during the General Debate, with much of the discussion around AI's "[potential to either facilitate development or perpetuate inequality](#)."

Other multilateral frameworks that must proactively center conflict prevention include the OECD's [AI Principles](#), which aim to promote innovative, trustworthy AI that respects human rights and democratic values and should be commended for being ahead of its time. The World Bank uses [AI in risk and crisis management](#), such as examining satellite images to identify humanitarian vulnerabilities or analyzing weather and earthquake stations to predict potential natural disaster impact. While these use cases are a start, the World Bank still needs to institutionalize the peaceful use of AI within any organizational strategy, especially those around peace and security.

Policymakers and the international community must more explicitly articulate and address AI's direct connection to global peace and security. Bilateral governments, the private sector, civil society, and multilateral institutions must take urgent action to mitigate the potential impacts of AI and work to ensure it promotes social cohesion and peacebuilding and minimizes its adverse effects on peace and security globally.

Multilateral institutions can critically fill gaps between national efforts to ensure a coordinated international effort that provides AI technologies to uphold peace and security. Multilateral frameworks can outline principles and best practices for using AI technologies to prevent violence and atrocities and promote peace, which can help rally good actors and civil society for a common cause. The UN should also consider and integrate AI into its [New Agenda for Peace](#)—a vision

that outlines how multilateral efforts for peace and security can be revamped given the world's rapid transition—and [Summit of the Future](#) in 2024.

Regulation by governments should prohibit the use of AI to create content and disinformation at a wide scale that supports violence, calls for attacks based on gender, race, ethnicity, or other forms of identity,

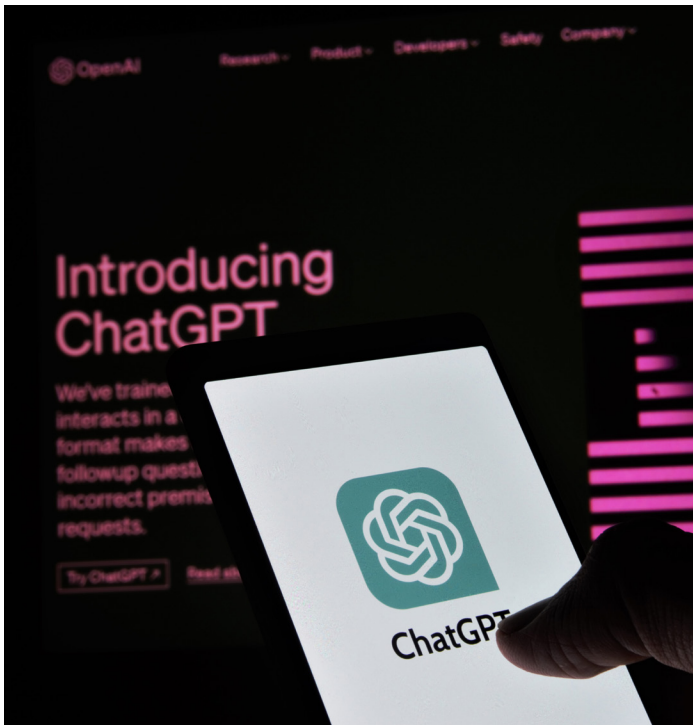
undermines human rights, and urges violations of human rights and atrocities. Given that Section 230 [does not protect creators or developers of online content](#), regulation should ensure liability for companies that design generative AI tools that then contribute to violence, atrocities, and human rights violations. Regulation could also require that generative AI tools include watermarks or computer code that identifies their AI-generated content to ease the debunking of AI-produced disinformation. For instance, Adobe [stated](#) that its generative Firefly tool will feature "nutrition labels"—including the date of

an image's creation and the digital tools used to create it—as it integrates with Google's Bard AI chatbot. Verifiable but refined methods of watermarking AI-generated content—[such as tweaking every 10th image pixel or using unique coding identifiers](#)—can help withstand alterations standard to disinformation, such as cropping and resizing. These tools can then better debunk disinformation on social media and prevent manipulation.

Data privacy regulation should also ban AI programs, such as facial recognition or image/audio generation technology, from being trained on personally identifiable information pulled from the internet, such as images posted on social media, voices, and names. All actors running confidential databases that utilize AI, such as law enforcement and government records, must take comprehensive steps to safeguard personally identifiable information through data risk assessment and protection, including data encryption and robust cybersecurity protocols.

Governments could also prohibit or issue licenses that

Multilateral frameworks can outline principles and best practices for using AI technologies to prevent violence and atrocities and promote peace, which can help rally good actors and civil society for a common cause.



[Mojahid Mottakin/Unsplash]

allow exclusive use of the most powerful forms of AI technologies, similar to the EU's proposed AI Act. These [risk categories](#) would prohibit AI tools that present "unacceptable risk," such as social scoring systems and automated monitoring of people in public spaces. Limited licenses for high-risk but not prohibited forms of AI, like critical infrastructure or medical devices, should ensure that AI programs with the potential to impact the security, freedom, and rights of people have safeguards in place so that the personally identifiable information of citizens is not publicly shared and that citizens' livelihoods are not placed under AI or autonomous decision-making. In ranking the potential risks of AI technologies, technologists must consider the impact of a technology if incorrect decisions are made and how the technology could decrease security, freedom of expression/privacy, or other civic rights in its use.

Given valid concerns that passing regulation in the U.S. or EU could unequally focus technology company resources in those more developed countries as they adhere to the new legislation—and thus [divert them away from regulatory and safeguarding efforts in other, more fragile countries](#)—regulation must also require "investments that platforms make on content moder-

ation and how these resources are deployed across the world." These provisions could include requiring a minimum monitoring presence and preventative interventions in countries with high user bases and severe disinformation risk.

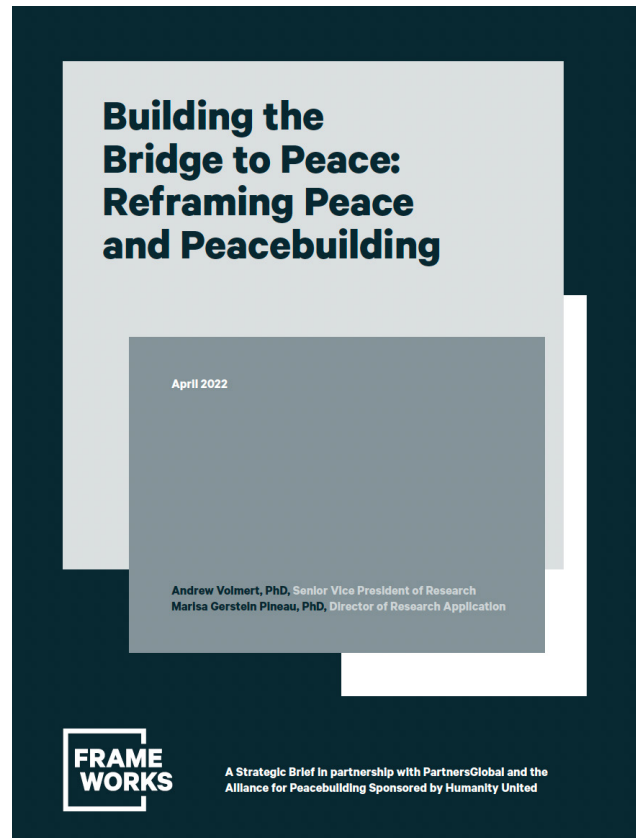
An agency that oversees these safeguards that protect citizens' security, freedom, and rights should lead and monitor AI regulation. To identify potential risks and prevent harm, the agency should regularly assess technology companies and stakeholders that design and deploy AI tools. It should also include representation from senior government officials overseeing U.S. foreign assistance, particularly conflict and atrocity prevention, to ensure that enforcement of AI regulation dovetails with global peacebuilding efforts. Those who violate AI laws and regulations should face hefty fines and incarceration.

In the U.S., Congress can require the Administration to create a whole-of-government AI strategy that touches foreign and domestic policy. Foreign assistance components of the strategy would include diplomacy, development, and defense agencies—as well as some other foreign-facing agencies like the Department of Treasury—and aim to promote peace and security abroad through mitigation of AI's negative impacts and innovative use of AI to build peace and prevent conflict. The domestic component of the strategy could include the Department of Justice and the Department of Homeland Security and focus on addressing AI-related drivers of domestic extremism and violent conflict, as well as the peaceful use of AI in those efforts. Like the [Global Fragility Act](#) and the [Elie Wiesel Genocide and Atrocities Prevention Act](#), the strategy could include regular reporting to Congress to allow the public and civil society to hold the government accountable.

2. Develop AI-powered tools that “translate” divisive rhetoric into language that prevents and reduces violent conflict, violence, and fragility and builds sustainable peace.

To counter AI’s potential to drive disinformation and online/offline violence, the peacebuilding field, AI engineers, and the private sector must collaborate to develop tools that amplify narratives proven to increase understanding and support for peacebuilding. For example, peacebuilders could work with AI engineers, particularly those skilled in [natural language processing](#), to develop AI-powered language tools that detect divisive and manipulative rhetoric on social media characteristic of disinformation and other online violent speech and then allow people to “translate” that content into language that promotes peace in online and offline conversations.

While disinformation facilitates offline violence and worsens trust and social cohesion, the Alliance for Peacebuilding’s [peace framing and narrative research](#) found that the use of certain narratives can increase understanding of and support for peacebuilding approaches. These narratives include portraying peace as an active and ongoing process (like building a bridge) rather than a passive end state, as well as showing that we are all interconnected and therefore what happens in one place in the world impacts us all. AI-powered tools that amplify and encourage these narratives could shift online conversations from divisive rhetoric to promoting and building peace online and offline.



Building Good Narratives: Examples

Since 2001, the U.S. military has intervened internationally over and over under the auspices of the “War on Terror.” Using the military to address violent conflict has exacerbated sectarian tensions and led to widespread violence. Instead, we must use the proven tools of peacebuilding.



The pandemic has underlined just how *interconnected* the world is—what affects one part of the world *affects all of us*. When we allow violent conflict to happen anywhere, it can *spread and disrupt peace everywhere*. As a global community, peacebuilding matters to all of us.

Peacebuilding programs are designed to end hostilities and restore peace. When violent conflict erupts, peacebuilders stop the violence by bringing together the parties to resolve conflict without violence. Once the conflict is resolved, peacebuilding programs re-establish safety and stability.



Peacebuilding programs work to create *sustained* peace in countries and communities around the world. These programs engage in *ongoing efforts* to address the root causes of violent conflict. By *building and maintaining* the conditions for peace, peacebuilders work to ensure *continuing* safety and stability.



Alliance for Peacebuilding

About Alliance for Peacebuilding:

Named the “number one influencer and change agent” among peacebuilding institutions worldwide—AfP is a nonprofit and nonpartisan network of 200+ organizations working in 181 countries to prevent conflict, reduce violence, improve lives, and build sustainable peace. At our core, AfP cultivates a network to strengthen and advance the peacebuilding field, enabling peacebuilding organizations to achieve greater impact—tackling issues too large for any one organization to address alone.

Cover Image: The Ars Electronica Garden Berlin: “Artificial Reality – Virtual Intelligence” by University of Applied Science Berlin – School of Culture and Design, Department of Communication Design (DE). [[Franziska Baur/Ars Electronica](#)]

Published by:

Alliance for Peacebuilding
1800 Massachusetts Ave, NW, Suite 401
Washington, DC 20036 USA
www.allianceforpeacebuilding.org

©October 2023

For additional information, please reach out to:

Nick Zuroski
Manager - Policy & Advocacy
nick@allianceforpeacebuilding.org