A Seat at the Table?

Artificial Intelligence and the Political Foundations of International Conflict Mediation

Author

Dr. Tanja Flanagan

Council on International Mediation
مجلس الوساطة الدولية

Council on International Mediation (CIM)

February 2026

A Seat at the Table? – Artificial Intelligence and the Political Foundations of International Conflict Mediation

Artificial intelligence has become embedded in the informational infrastructure of international conflict mediation. It rarely appears as a formal participant in negotiations, yet it increasingly shapes what mediators see and how conflicts are interpreted, what they prioritize, how risks are assessed, and what mediators treat as plausible. This development raises questions that extend well beyond efficiency. Mediation depends on consent, trust, and political ownership of outcomes. Artificial intelligence alters the conditions under which facts are produced, interpreted, and contested. Artificial intelligence accelerates analysis and expands information intake, but, crucially, it projects a degree of technical authority that can shift bargaining power. The central question is therefore not whether artificial intelligence improves mediation in a technical or efficiency sense, but whether its use strengthens or weakens the political foundations that make mediation possible.

A second question follows. Even if artificial intelligence helps mediators process information, does it change the substance of negotiations by narrowing the range of acceptable arguments, privileging quantifiable claims, and converting political disagreements into a technical dispute? This risk becomes clearer in resource negotiations, especially international water diplomacy, where data models often stand in for power. A 2025 research study published in Ambio – based on expert interviews – finds that artificial intelligence tools can shape knowledge production and exchange in water diplomacy, and that these tools constitute something of a double-edged sword: artificial intelligence tools can on the one hand support consensus building, but on the other hand reinforce inequalities through opacity, access gaps, and a false appearance of neutrality.

These risks become more urgent as AI-generated content proliferates across the global information environment. The ability of artificial intelligence to generate plausible narratives, synthetic evidence, and persuasive analytical outputs intensifies long-standing challenges related to misinformation, manipulation, and epistemic authority. Mediators already operate in environments saturated with contested claims. AI increases the volume, speed, and sophistication of those claims, raising the stakes of how information is filtered, validated, and presented. Oversight, transparency, and accountability are no longer ethical add-ons; rather, they are functional requirements for mediation credibility.

This essay analyzes how artificial intelligence is currently used in international conflict mediation, which stages of the mediation lifecycle it most strongly affects, the advantages it offers, and the risks it introduces. It draws on practice within multilateral diplomacy, preventive diplomacy, and post-conflict implementation, with particular attention to international water diplomacy as a revealing case. It then maps these findings onto a mediation lifecycle and specifies design rules and red lines that mediators should enforce. The essay concludes with a normative, effective, and substantive assessment of whether and under what conditions artificial intelligence belongs in international conflict mediation.

Section I. What counts as artificial intelligence in mediation practice?

In mediation contexts, artificial intelligence refers less to autonomous decision-making systems than to a family of analytical and support tools that reshape information environments. In this field, it covers at least five tool families.

First, automated language processing, including translation, transcription, summarization, topic clustering, and entity extraction. These tools sit at the core of modern information management.

Secondly, social media and open source analytics, including classification of posts, detection of narrative spikes, network mapping of online actors, and triage of misinformation signals.

Third, predictive and forecasting systems, ranging from basic risk scoring to more complex models trained on historical patterns.

Fourth, decision support for drafting and process work, including clause comparison, option generation, and consistency checks across texts.

Fifth, domain modeling systems in climate and resource diplomacy, including hydrological models, drought and flood forecasting, and infrastructure impact simulation.

None of these AI systems mediates in the human sense. They do not negotiate tradeoffs. They do not build trust. They do not read a room. They do not accept responsibility. But they do influence mediation indirectly by structuring attention, framing uncertainty, and influencing which claims appear credible or actionable. Artificial intelligence tools shape mediation by shaping information, and information in turn shapes bargaining. This distinction is essential. When artificial intelligence is treated as political infrastructure, it can support mediation because political contestability and legitimacy is preserved. When artificial intelligence is treated as political authority, though, it undermines mediation by displacing contestation with technical certainty. Mediation is, after all, a political facilitation, not a technical arbitration.

Section II. Does it make sense to use artificial intelligence in international conflict mediation?

Artificial intelligence makes sense in international conflict mediation only if it is treated as a form of decision support for mediators rather than as a substitute for mediation itself. Its legitimate role lies in supporting human judgment, not replacing it. In practice, this role clusters around four parts of the mediation cycle.

First, artificial intelligence supports mediators in the preparation cycle by organizing, translating, and summarizing large volumes of text, media, and stakeholder input. This function helps mediation teams cope with information overload and reduces delays between signal and response.

Second, artificial intelligence supports process design and inclusion by assisting in stakeholder mapping, widening consultation, and testing participation options for groups excluded by

security constraints, geography, or political barriers. This use is valuable but politically sensitive, because inclusion decisions shape legitimacy.

Third, artificial intelligence supports situational awareness during talks by tracking narratives, misinformation, and escalation indicators in near real time; then feed structured briefs to a mediation team. This allows mediation teams to anticipate external shocks that could destabilize negotiations.

Fourth, language support: Translation and multilingual summarization reduce friction in multi party and multilingual talks. This can widen participation and improve access to non English sources, which improves substantive understanding.

Fifth, artificial intelligence can support implementation by helping monitor commitments, detect signals of compliance or backsliding, and prioritize necessary follow-ups.

This does not mean artificial intelligence belongs everywhere. Artificial intelligence does not make sense when it enters negotiations as an unchallengeable object, when it claims unquestionable neutrality, or when it substitutes technical authority for trust building. The Ambio water diplomacy findings underscore this point. In water negotiations, models often stand in for power. Artificial intelligence intensifies this effect by shaping what counts as valid knowledge and credible data. Without careful governance ensuring equal access and understanding, this undermines confidence rather than building it.

Section III. How artificial intelligence has been used in conflict mediation

Artificial intelligence has entered mediation practice primarily through institutions and organizations that are responsible for political analysis and mediation support. Artificial intelligence is primarily used as a backstage analytical tool within mediation teams rather than as a visible authority in negotiations. AI systems support political analysts, mediation support units, and advisers by processing large volumes of data, identifying narrative shifts, summarizing reports, and modeling scenarios.

Their outputs inform the mediator’s understanding and preparation, but they are typically filtered, interpreted, and translated into political language before anything is shared with the negotiating parties. The systems themselves do not usually ‘speak’ to principals such as heads of state, ministers, or armed group leaders, nor are their outputs presented as binding evidence or neutral determinations of fact. This internal use preserves mediator discretion and legitimacy. If AI outputs were introduced directly to principals as authoritative or objective judgments, they could alter bargaining dynamics, trigger disputes over model validity, and undermine trust by appearing to replace political judgment with technical authority.

One prominent use is social media and information environment scanning for early warning and political analysis. The United Nations Department of Political and Peacebuilding Affairs has used a social media scanning tool known as Sparrow to help analysts separate meaningful

political discourse from noise. The tool supports early warning and live monitoring by identifying narrative spikes, coordinated activity, and shifts in online engagement. Public UN materials describe Sparrow as a tool to help analysts and desk officers separate noise from authentic political speech and to strengthen early warning and live monitoring of unfolding crises. Importantly, Sparrow informs attention and prioritization. It does not generate determinations of intent or legitimacy. It does not decide policy. What it does is support attention allocation.

A related use of AI appears in ‘e-analytics’ support to peacemaking and preventive diplomacy. UN departments and partners have produced guidance and training on using data analytics, natural language processing, and machine learning to support political analysis and peacemaking support. This work consistently frames technology as a support to human judgment rather than a replacement for it. Technology is treated as a set of methods to support analysis rather than as a substitute for political work and negotiations.

Data mining and visualization tools have also been developed to support peace and security analysis. UN Global Pulse has experimented with extracting and analyzing data from public sources such as social media and radio broadcasts, including tools designed to detect rumor dynamics and social tension. The Global Pulse report on experimenting with big data and artificial intelligence describes a tool named QataLog intended to support extraction and analysis of public discussions, including radio, for peace and security related work. This matters for mediation because narrative monitoring and rumor detection affect conflict dynamics. Public discourse shapes mobilization, legitimacy, and elite signaling. AI assisted analysis can help mediators anticipate external shocks that affect bargaining space.

Mediator-facing practice guidance has expanded alongside these tools. Platforms such as the UN Peacemaker initiative host thematic guidance on how digital technologies are transforming politics and conflict, emphasizing adaptation without abandoning core mediation principles. Negotiation research and practitioner discussion increasingly describe AI as a coach or advisor for negotiators. In mediation contexts, this shows up as training support, rehearsal, and option exploration rather than operational decision making in live talks. This direction aligns with the idea of ‘AI as copilots’ that support preparation and drafting, while humans retain adjudication, and judgment in the negotiating rounds.

Academic work and policy research explore machine learning approaches tied to conflict dynamics and mediation related outcomes. These efforts focus on improving assessment and triage rather than automating negotiation. Similarly, negotiation research increasingly explores artificial intelligence as a coaching or rehearsal tool, supporting preparation and training rather than live bargaining.

International water diplomacy presents a distinct and revealing pattern. Here, artificial intelligence does not remain in the background. Hydrological models, forecasting tools, and scenario simulations have increasingly entered agenda setting, framing, and justification. Research published in *Ambio*, based on interviews with practitioners, shows that AI-driven tools influence how knowledge is produced and negotiated in transboundary water governance. States introduce models to support their claims. These tools can support weaker actors by

lowering technical barriers, but it can also reinforce inequalities when access, transparency, and literacy diverge. AI models can shift bargaining power because parties with stronger tools can define the baseline that others must argue against.

The Ambio study and the Swedish Defence University account provide a sharper case because water diplomacy ties negotiation outcomes to data modeling and forecasts. The study draws on expert interviews and examines how AI tools influence knowledge production and exchange in water diplomacy. It reports both optimistic views about improved objectivity and social learning, and concerns about power asymmetries, transparency, and capacity gaps. This logic scales to other mediation settings that involve territory, resources, or security guarantees.

Section IV. Areas of mediation affected by artificial intelligence

Artificial intelligence affects mediation unevenly across stages. Its influence is strongest before formal talks begin and after agreements are signed. Its role during direct bargaining remains narrow and contested.

Early Warning and Preventive Diplomacy

Early warning represents the most established and least controversial application. Artificial intelligence can help identify correlations between rhetoric, mobilization, and assess rising risk across time and geography. UN documentation and related research describe the use of social listening tools such as Sparrow to expand monitoring across political contexts. Analysts receive structured outputs that highlight where attention is needed. Preventive engagement often follows quietly through informal channels.

In preventive diplomacy, timing matters. Earlier engagement often preserves leverage and expands room for compromise. AI helps reduce the chance that a small mediation team misses signals that precede escalation. Thus, this use raises fewer legitimacy concerns because outputs remain internal and advisory. Parties do not negotiate over model results. Mediators retain full discretion over interpretation and response.

Preparation, Conflict Mapping, and Mandate Design

Artificial intelligence shapes how conflicts are defined. It can support actor identification, relationship mapping, and problem structuring. Especially in resource disputes, AI-driven models increasingly define scarcity narratives and future risk scenarios.

This creates a risk of mandate capture. The risk of mandate capture arises when analytical models move from informing mediation to defining it. In many contemporary disputes, especially those involving territory, resources, climate stress, or infrastructure, artificial intelligence–driven models generate forecasts, quantify scarcity, and simulate future scenarios. These outputs can appear technically rigorous and internally coherent. When mediators rely heavily on such models to structure the negotiation mandate, the problem space becomes bounded by what the

model measures and projects. Issues that cannot be easily quantified, such as historical injustice and grievances, identity-based claims, colonial legacies, or perceived dignity and recognition, may fall outside the frame. As a result, the mandate narrows to what can be modeled rather than what is politically contested.

Technical coherence can therefore come at the expense of political contestability. A model might present a drought projection, river flow allocation, or demographic forecast as a neutral baseline. Yet that baseline embeds assumptions about data sources, time horizons, and normative priorities. If those assumptions go unexamined, they can silently privilege certain distributive outcomes while marginalizing others. Parties may find themselves arguing within the parameters of a technical simulation rather than over the underlying values and interests at stake. In this way, the authority of the model can crowd out political arguments, converting what is fundamentally a distributive and normative dispute into a debate over technical validity.

Process Design and Inclusion

AI-assisted stakeholder mapping can help mediators widen consultation as it can surface informal brokers and overlooked intermediaries. This can improve inclusion in fragmented conflicts. At the same time, digital visibility skews representation. Actors with limited online presence or operating in non-dominant languages risk marginalization. This creates risks of exclusion masked as efficiency. Given that the legitimacy of a mediation process depends heavily on perceived inclusion and fair representation, this is worrisome.

Agenda Setting and Problem Framing

Agenda setting is the most politically sensitive point of AI use. Agenda setting determines what is “on the table” and what is treated as background fact. When artificial intelligence enters at this stage, especially in resource negotiations such as transboundary water disputes, it does more than provide information; AI may increasingly define which futures appear negotiable.

In water diplomacy, hydrological models, climate projections, and infrastructure simulations can establish baseline assumptions about scarcity, future flows, or risk distribution. Once these projections are presented as the objective starting point, the scope of negotiation narrows to adjustments within the modeled parameters. Thus, as previously stated, political questions about equity, historical allocation patterns, development rights, or environmental justice can become secondary to debates over model calibration, data inputs, or forecast accuracy. What begins as a distributive and normative dispute risks being reframed as a technical disagreement about numbers. Water diplomacy thus reveals the deeper theoretical issue. Knowledge in negotiation is never purely technical. It is social and political.

The Swedish Defence University research highlights the trust implications of this shift. When one party introduces AI-generated forecasts and presents them as neutral or scientifically authoritative, other parties may perceive an imbalance, particularly if they lack access to the underlying model or the expertise to interrogate its assumptions. Even if the model is methodologically sound, the perception of asymmetry can erode confidence in the process.

Trust depends not only on the quality of data but on shared understanding and the ability to contest interpretations. If artificial intelligence becomes the arbiter of which futures are plausible or acceptable, political negotiation risks being displaced by technical validation, and the mediation process itself can lose legitimacy.

Negotiation Support and Drafting

Direct use of artificial intelligence during live bargaining remains narrow and largely procedural. Mediation teams may use AI tools to compare draft clauses across prior agreements, generate structured summaries of positions, or produce first-pass text for ceasefire provisions or implementation timelines. In these cases, human mediators and legal advisers retain full control. They review, revise, and approve every formulation before it enters the negotiating room. AI does not independently propose settlements or evaluate concessions. Its role is supportive and administrative rather than decisional.

The risk, however, lies in framing effects. Even when AI-generated language appears neutral, it is shaped by patterns in prior agreements and dominant legal or diplomatic practices embedded in its training data. This can subtly privilege certain institutional arrangements, sequencing logic, or definitions of compliance. For example, standard formulations drawn from past peace agreements may reflect particular power distributions or governance models that are not politically neutral in the present case. When negotiations begin from AI-assisted draft text, the starting point of discussion may already reflect assumptions about what is “normal,” “reasonable,” or “standard.” In this way, the technology does not dictate outcomes, but it can shape the baseline from which bargaining proceeds.

Implementation, Monitoring, and Compliance

Post-agreement monitoring has become one of the most attractive areas for artificial intelligence use because it promises more timely and independent detection of compliance problems. AI systems can analyze satellite imagery to flag unusual troop movements near buffer zones, process large volumes of local reporting to identify patterns consistent with ceasefire violations, or scan online discourse for inflammatory narratives that signal mobilization or incitement. In a resource-sharing agreement, AI could monitor water flow data against agreed allocation thresholds, or detect infrastructure disruptions that suggest non-compliance. These capabilities reduce reliance on self-reporting by parties and can help mediators intervene early before minor breaches escalate into renewed violence.

At the same time, this expanded monitoring capacity raises sovereignty and privacy concerns. If satellite imagery is supplied by a private company headquartered in one of the mediating states, a party may suspect political bias. If narrative tracking relies on scraping social media platforms that are unevenly used across communities, one side may claim selective surveillance. Even technical data, such as water flow sensors installed upstream, may be perceived as instruments of external control if they are funded or managed by actors aligned with one party. In fragile post-conflict environments, perceptions matter as much as accuracy. Monitoring systems that

appear intrusive, externally controlled, or asymmetrically applied can undermine confidence in the agreement they are meant to protect.

Section VI. Disadvantages and risks of artificial intelligence in conflict mediation

The risks of artificial intelligence in mediation are structural rather than incidental; and the risks matter more than the benefits if mediators treat AI as an authority rather than a means of support. The following areas seem to hold the highest risks.

Bias and uneven error rates: Training data skews toward high visibility actors and dominant languages. This can marginalize communities already excluded. Even when outputs look neutral, the data pipeline often encodes power.

Opacity and contestability problems: Parties may reject AI informed claims if they cannot inspect methods, data, and assumptions. Proprietary tools make this worse. Contestability sits at the heart of legitimate mediation. A black box undermines that legitimacy. Negotiators need to have the ability to question data sources and assumptions and to understand how tools shape what becomes possible.

Misinformation intake risk: Social platforms amplify bots, propaganda, and hate speech. Bad inputs produce bad outputs unless mediation teams apply strict validation. Tools like Sparrow explicitly aim to separate noise from authentic political speech, which signals how serious this risk is in practice.

Manipulation incentives: Parties may try to game models by flooding channels with coordinated content or staged signals. If mediators treat model outputs as evidence without questioning sources or training data, actors gain incentive to shape the evidence.

Confidentiality and security: Mediation depends on trust. AI tools introduce data handling, vendor access, and leakage risks, especially with cloud services. Even internal use can create exposure if teams upload sensitive notes or draft texts to systems outside secure infrastructure.

Overreliance and deskilling: Teams may defer to model outputs and underweigh local knowledge, field reporting, and intuition built from relationships and experience. Mediation work rewards contextual judgment. Overreliance on AI erodes that capacity.

Legitimacy and accountability: If an AI supported recommendation leads to harm, responsibility remains with the mediator, yet decision chains get harder to audit. This creates both an ethical risk and a political risk. Legitimacy and accountability become harder to trace. Responsibility remains human, but decision chains grow opaque.

Political backlash: Some actors frame AI as foreign surveillance or as biased Western technology. This can erode party buy-in and trigger a refusal to engage; and the willingness to

enter and engage faithfully in a mediation process, is crucial. This risk rises when AI tools rely on external vendors or donor funded systems.

It should be emphasized that AI is not inherently good or bad. Its effects depend on context and governance. Experts stress that AI should be one tool among many, enhancing rather than displacing human-led diplomacy. The key question is who decides what counts as valid knowledge and how that decision shapes power. This insight generalizes beyond water. All mediation depends on contested facts. Artificial intelligence reshapes epistemic authority. Mediators must therefore govern not only outputs but perceptions of legitimacy.

Section VII. Artificial intelligence and international diplomacy as a strategic environment

Mediation unfolds within a broader diplomatic ecosystem that shapes how technology is perceived and contested. Artificial intelligence is no longer a narrow technical innovation. It operates across innovation policy, regulation, military doctrine, surveillance practices, communication ecosystems, and even cultural production. Because AI influences economic competitiveness, strategic autonomy, and domestic governance, it carries political meaning before it ever enters a mediation room. Parties do not encounter AI as a blank tool. They encounter it through the lens of their own experiences with digital power, regulatory debates, and geopolitical rivalry. This context shapes expectations about sovereignty, fairness, and legitimacy.

States also compete over the rules and standards that govern artificial intelligence. Debates about export controls, data governance, surveillance norms, and algorithmic accountability increasingly form part of global diplomacy. Artificial intelligence also reshapes diplomacy itself. Analysis by the Konrad-Adenauer-Stiftung frames AI as a strategic factor comparable to energy security or standards-setting. Thus, states compete over narrative authority and information advantage. AI enhances a state's capacity to generate persuasive content, analyze public opinion, and shape discourse at scale. These dynamics affect mediation indirectly. A state that views AI as an extension of strategic influence may interpret AI-supported analysis in a negotiation as a form of leverage rather than neutral support. Perception, in such contexts, matters as much as methodology.

For this reason, even analytical tools presented as neutral can be interpreted as instruments of power. If one party controls the data infrastructure, the vendor relationships, or the modeling expertise, AI-assisted outputs may appear aligned with that party's interests regardless of intent. Guardrailing the interaction between politics and AI therefore becomes essential. Transparency about assumptions, shared access to tools, clear lines of human accountability, and open acknowledgment of uncertainty are not procedural niceties. They are prerequisites for maintaining trust in a diplomatic environment in which technology itself has become a contested strategic domain.

Thus, the global strategic context should reinforce three practical lessons for mediators.

First, AI expands capacity for analysis, forecasting, and narrative production. This accelerates decision cycles and rewards actors with stronger digital infrastructure.

Second, AI becomes a subject of diplomacy itself through debates about regulation, standards, surveillance, and military use. These debates shape trust between parties in mediation contexts.

Third, AI reshapes normative competition. Competing governance models influence how parties interpret AI use in mediation. Some actors will treat AI as an extension of state surveillance. Others will treat it as a neutral modern tool. Mediators must anticipate these perceptions.

The key implication is that legitimacy will not follow from technical sophistication. Legitimacy will follow from transparency, shared access, and clear human accountability.

Section VIII. AI red lines in the mediation cycle

A mediation lifecycle runs from pre engagement through agreement implementation. Artificial intelligence affects each stage differently. Mediators are thus required to adopt stage specific design rules and enforceable red lines that protect the legitimacy of the mediation process and political ownership.

As previously stated, at the pre-mediation stage, artificial intelligence plays its most legitimate and least controversial role. Nevertheless, AI tools should remain in a subordinate position, its uses limited to advisory and internal processes. AI outputs should guide attention and prioritization. They should not generate conclusions presented as authoritative assessments of intent, blame, or legitimacy. Mediators should not share AI outputs with parties as determinations of culpability or likelihood of violence. Once outputs acquire external authority, they begin shaping political narratives in ways parties cannot contest. Mediators should also avoid a narrow data monoculture. Systems should rely on diverse data sources. Exclusive reliance on a single platform, dominant language streams, or donor curated datasets entrenches asymmetries and distorts signals.

During the conflict assessment and mandate design stage, AI tools tend to support conflict mapping, actor identification, and problem structuring. In resource disputes, AI driven models can define the boundaries of the issue space through scarcity scenarios and future risks. Any AI supported assessment used to inform mandate design should disclose assumptions, data gaps, and uncertainty ranges. The mediator’s task is not technical coherence alone. The mediator’s task is preserving political contestability. This means that AI generated models should not capture mandates. Mediators should not allow models to define the scope of negotiations in ways that exclude distributive claims, historical grievances, or normative arguments. If a model implies that certain futures are non negotiable, it exceeds its role. Mediators must retain ownership of the mandate rationale. If technical models become the justification for engagement

terms, responsibility shifts away from political judgment and undermines the legitimacy of the whole process.

During the process design and inclusion stage, any AI assisted mapping that informs inclusion should be transparent to all parties, or balanced through parallel outreach channels. Mediators should err on the side of caution, and should assume digital visibility correlates imperfectly with political relevance. Meaning, mediators should not exclude or deprioritize actors based solely on AI generated assessments that affected actors cannot interrogate. And AI tools are not a substitution for field engagement. AI tools should complement human engagement. If digital analysis substitutes for relational work, the process loses legitimacy even if speed and efficiency increases.

Stage 4, the agenda setting stage, is the point of highest sensitivity. AI generated models should be treated as one lens among several. Mediators should invite alternative interpretations, counter scenarios, and non technical claims. Mediators should avoid any neutrality claims that foreclose debate. Meaning, mediators should not permit any party to introduce AI outputs as politically neutral evidence that ends discussion. Once technical outputs become neutral facts, negotiation shifts into technical arbitration. In addition, if one party controls proprietary models or exclusive data, mediators should restrict their role or provide compensatory support to other parties. Allowing unequal technical authority at the agenda stage distorts outcomes early and often irreversibly.

During the direct negotiation and bargaining stage, AI should have its narrowest legitimate role. Its use should focus on internal support functions such as drafting assistance, comparative text analysis, and scenario visualization requested by all parties. All proposals, tradeoffs, and formulations should be explicitly attributed to human judgment. AI outputs can inform drafts. They should not appear as recommendations in their own right. Mediators should not allow AI systems to suggest optimal compromises, fair distributions, or equilibrium outcomes during bargaining. Settlement requires ambiguity, creativity, and political ownership. Algorithmic optimization undermines these features. Finally, AI tools should not be invoked in ways that signal inevitability or superior rationality. Visual power can coerce even without intent. Mediators must manage that risk.

In the agreement drafting and validation stage, every AI assisted draft element should remain traceable to human decision making. Mediators and parties should explain why language appears as it does without reference to algorithmic authority. If parties cannot distinguish between politically agreed language and system generated text, accountability erodes. All agreements must remain human artifacts.

During the final stage of conflict mediation, implementation, monitoring, and compliance, any AI monitoring tools in use should match the sensitivity of the agreement and respect the sovereignty, privacy, and consent of the parties involved. Monitoring should support confidence building, not be an enforcement by exposure. In this vein, there should be no unilateral surveillance. Mediators should not deploy monitoring systems that one party perceives as externalized control or intelligence collection. Monitoring legitimacy depends on a previously

agreed upon shared scope. Finally, AI detected violations should never trigger action without human review. Automated inference must not replace political judgment in fragile environments.

Across all stages, one rule dominates. Artificial intelligence should function as infrastructure rather than authority. AI shapes information environments, not outcomes. It informs judgment, not legitimacy. AI is merely political infrastructure, not a political actor. AI must not substitute for trust. Once technical authority replaces relational legitimacy, mediation stops functioning as mediation.

Conclusion

It makes sense to use artificial intelligence in international conflict mediation under strict constraints. It makes sense when mediators use it to widen information intake, speed synthesis, strengthen early warning, and support implementation monitoring. It makes sense when it helps weaker actors gain analytical capacity through open access tools and shared modeling processes. It makes sense when it exposes assumptions rather than masking them.

It does not make sense when it concentrates expertise, claims neutrality, displaces trust with technical authority, or becomes a gatekeeper for participation. It does not make sense when proprietary tools and opaque models define the agenda and narrow the bargaining space in ways parties cannot contest. It does not make sense when mediators delegate political judgment to systems that cannot accept responsibility.

The practical bottom line is discipline. Keep humans responsible for judgment, relationship building, and final decisions. Require transparent methods, secure data handling, and active checks for manipulation and bias. Treat AI outputs as contested inputs, not neutral facts. Use red lines to preserve legitimacy at each stage.

International water diplomacy offers both early warning and proof of concept. AI can help create shared baselines and improve social learning. It can also deepen politicization and undermine trust when access and understanding diverge. The same pattern holds across international conflict mediation. Governance determines whether AI supports peace or hardens inequality.

Source List

Department of Political and Peacebuilding Affairs (DPPA). *"We Use Technology, Not the Other Way Around: Social Media and Political Analysis."* United Nations DPPA Medium, 2019.

United Nations Department of Political and Peacebuilding Affairs (DPPA). *E-Analytics for Peacemaking and Preventive Diplomacy.* United Nations, various editions.

United Nations Global Pulse. *Experimenting with Big Data and AI to Support Peace and Security.* New York: United Nations, 2019.

United Nations Peacemaker. *Digital Technologies and Mediation Toolkit.* United Nations Mediation Support Unit, various updates.

United Nations. *Strategy for the Digital Transformation of UN Peacekeeping.* New York: United Nations, 2021.

Stockholm International Peace Research Institute (SIPRI) and Swedish Defence University. *Artificial Intelligence in International Water Diplomacy.* Ambio (2025).

Swedish Defence University. *AI Reshapes International Water Diplomacy.* Institutional research summary, 2025.

Konrad-Adenauer-Stiftung. *Künstliche Intelligenz: Politik, Regulierung und strategische Implikationen.* Berlin: Konrad-Adenauer-Stiftung, 2024.

Stimson Center. *New Security Beat.* "Shall We Ask ChatGPT About Water During Negotiations?" November 17, 2025.

United Nations General Assembly and Security Council. *Our Common Agenda: Report of the Secretary-General.* A/75/982, 2021.

Harvard Program on Negotiation. *Artificial Intelligence and Negotiation Practice.* Cambridge, MA: Harvard Law School, various publications.

United Nations Institute for Disarmament Research (UNIDIR). *AI and International Peace and Security: Implications for Governance.* Geneva: UNIDIR, 2023.

Does it make sense to use AI in international conflict mediation?

Yes, if you treat AI as decision support for mediators, not as a substitute for mediation. AI fits best in four parts of the mediation cycle.

Preparation. Organize, translate, and summarize large volumes of text, media, and stakeholder input.

Process design and inclusion. Map stakeholders, widen consultation, and test participation options for groups excluded by security, distance, or politics.

Situational awareness during talks. Track narratives, misinformation, and escalation indicators in near real time, then feed structured briefs to a mediation team.

Implementation support. Help monitor commitments, verify signals of compliance or backsliding, and prioritize follow-up.

How has AI been used in conflict mediation?

1. Social media and information environment scanning for early warning and political analysis

   UN DPPA has used a dedicated social media scanning tool called Sparrow for early warning and analysis, built to help analysts separate signal from noise in online discourse.

2. “E-analytics” support to peacemaking and preventive diplomacy

   DPPA, UN Global Pulse, and partners have run training and produced guidance on using data, including machine learning and NLP methods, for peacemaking and preventive diplomacy.

3. Data mining and visualization tools to support peace and security analysis

UN Global Pulse developed QataLog to extract, analyze, and visualize data from sources such as social media and radio broadcasts, aimed at development, humanitarian, and peace efforts.

4. Digital toolkits and mediator-facing practice guidance

   The UN “Peacemaker” platform hosts thematic guidance on digital technologies, including pilots and tools used by peacemakers for analysis and process support.

5. Research prototypes for forecasting and mediation support

   Academic work has tested machine learning approaches tied to mediation and conflict process outcomes, aimed at improving assessment and decision support rather than replacing negotiators.

6. Training and applied experimentation in negotiation support

   Recent negotiation research and practitioner discussions describe AI as a coach or advisor for negotiators working in high-stakes contexts, focused on context-aware support.

Advantages of using AI in conflict mediation

Faster synthesis of messy information. AI helps teams process social media, local news, transcripts, and reports, then surface themes and shifts.

Earlier detection of escalation risk. Pattern detection helps analysts spot narrative spikes, hate speech trends, or mobilization signals sooner.

Better stakeholder mapping and inclusion. AI-assisted clustering and entity extraction helps map factions, influencers, and local intermediaries, which supports wider consultation.

More consistent drafting and “text work.” AI speeds up first drafts of agendas, options papers, talking points, and comparative clause language, then humans revise.

Language support. Translation and multilingual summarization reduce friction in multi-party and multi-lingual talks.

Scenario exploration. Models help teams test “if-then” pathways, identify tradeoffs, and stress-test sequencing, especially when time pressure is high.

Lower operational burden on small mediation teams. AI reduces repetitive work so senior staff spend more time on relationship management and judgment.

Disadvantages and risks in conflict mediation

Bias and uneven error rates. Training data skews toward high-visibility actors and dominant languages. This can marginalize communities already excluded.

Opacity and contestability problems. Parties may reject AI-informed claims if they cannot inspect methods, data, and assumptions.

Misinformation intake risk. Social platforms amplify bots, propaganda, and hate speech. Bad inputs produce bad outputs unless teams apply strict validation.

Manipulation incentives. Parties may try to game models by flooding channels with coordinated content or staged “signals.”

Confidentiality and security. Mediation depends on trust. AI tools introduce data handling, vendor access, and leakage risks, especially with cloud services.

Overreliance and deskilling. Teams may defer to model outputs and underweight local knowledge, field reporting, and intuition built from relationships.

Legitimacy and accountability. If an AI-supported recommendation leads to harm, responsibility still sits with the mediator, yet decision chains get harder to audit.

Political backlash. Some actors frame AI as foreign surveillance or as biased “Western tech,” which can erode buy-in.

How is AI currently used in mediation, and how is it planned to be used next?

Current use, what exists in real practice today

Analyst-facing tools for early warning and narrative tracking, like Sparrow, used by UN political teams for scanning and analysis.

Data mining and visualization for peace and security analysis, including QataLog and related e-analytics workflows.

Digital mediation practice guidance through UN mediation support channels, focused on practical integration of technology into mediation work.

Training and experimentation across the UN peace and security system on AI and digital methods for political analysis and peacemaking support.

Near-future direction, what programs and public plans point toward

More forecasting for specific conflict drivers. DPPA innovation reporting describes AI and machine learning work aimed at early warning forecasts for certain conflict types, including resource and mobility-linked conflict patterns such as transhumance-related events.

More integration with inclusion agendas. UN process work has begun linking AI-focused initiatives with women, peace, and security inclusion efforts in active files, signaling broader adoption in process design and participation support.

More mediator “copilots,” fewer standalone dashboards. Negotiation research and practitioner uptake point toward tools that support preparation, option generation, and context-specific coaching, embedded in daily mediator workflows.

More emphasis on governance. UN and policy research increasingly frames AI for peace as a risk-managed domain, with stronger focus on safeguards, transparency, and responsible use in fragile contexts.

A practical bottom line

Use AI to widen information intake, speed synthesis, and strengthen early warning.

Keep humans responsible for judgment, relationship building, and final decisions.

Require transparent methods, secure data handling, and “red team” checks for manipulation and bias.

Pre-mediation analysis and early warning

At the pre-mediation stage, artificial intelligence plays its most legitimate and least controversial role. Here, AI functions as analytical infrastructure that expands situational awareness without directly shaping political bargaining. Systems process large volumes of open source reporting, social media discourse, satellite data, and field reporting to identify escalation patterns, narrative shifts, and emerging actors. This use aligns with preventive diplomacy rather than adjudication.

The design rule at this stage is strict subordination. AI outputs must remain advisory and internal. They should guide attention and prioritization, not generate conclusions presented as authoritative assessments of intent, blame, or legitimacy. Mediators should treat AI as a triage instrument, not an oracle.

A first red line follows directly. AI outputs must never be shared with parties as determinations of risk, culpability, or likelihood of violence. Once model outputs acquire external authority, they stop supporting prevention and begin shaping political narratives in ways parties cannot contest.

A second red line concerns data scope. Systems must rely primarily on open and diverse data sources. Exclusive reliance on platform data, dominant languages, or donor curated datasets entrenches existing asymmetries and distorts early warning signals.

Conflict assessment and mandate design

Once mediation engagement becomes likely, artificial intelligence increasingly shapes how the conflict itself is defined. At this stage, AI supports conflict mapping, actor identification, and problem structuring. In water diplomacy and other resource disputes, AI driven models often begin to define the boundaries of the issue space by specifying scarcity scenarios, future risks, or technical constraints.

The design rule here is exposure of assumptions. Any AI supported assessment used to inform mandate design must be accompanied by explicit articulation of its underlying assumptions, data gaps, and uncertainty ranges. The mediator's task is not to present a technically coherent picture, but to ensure political contestability remains intact.

The central red line at this stage is mandate capture. Mediators should not allow AI models to define the scope of negotiations in ways that exclude distributive claims, historical grievances, or normative arguments. If a model implies that certain futures are "non negotiable," it has exceeded its legitimate role.

A secondary red line involves authorship. Mediators must retain ownership of the mandate rationale. If technical models become the justification for engagement terms, responsibility shifts away from political judgment and undermines mediator legitimacy.

Process design and inclusion

During process design, artificial intelligence affects who participates, whose knowledge counts, and which voices gain visibility. Stakeholder mapping tools, network analysis, and discourse clustering increasingly influence invitations, consultation formats, and sequencing decisions. In fragmented conflicts, this capacity offers real value. It reveals informal brokers and overlooked intermediaries.

The governing design rule at this stage is symmetry of access. Any AI assisted mapping or categorization that informs inclusion decisions must be transparent to all parties or balanced through alternative consultation channels. Mediators should assume that digital visibility correlates imperfectly with political relevance.

The red line here concerns exclusion by opacity. No actor should be excluded or deprioritized based solely on AI generated assessments they cannot interrogate or understand. When AI becomes a gatekeeper for participation, it reproduces inequality under the guise of efficiency.

A second red line concerns substitution. AI tools should complement, not replace, field engagement and human judgment. If digital analysis begins to substitute for relational work, the mediation process loses legitimacy even if it gains speed.

Agenda setting and problem framing

Agenda setting marks the point at which artificial intelligence becomes most politically sensitive. In water diplomacy, hydrological models, climate projections, and infrastructure simulations

increasingly define which issues appear discussable and which are framed as technical facts. At this stage, AI moves from background support to agenda shaping.

The central design rule is pluralization of frames. AI generated models should be presented as one possible lens among several, not as the definitive representation of reality. Mediators must actively counter the aura of objectivity by inviting alternative interpretations, counter scenarios, and non technical claims.

The primary red line is neutrality claims. Mediators must not permit any party to introduce AI outputs as politically neutral evidence that forecloses debate. Once technical outputs are treated as neutral facts rather than contested inputs, negotiation collapses into technical arbitration.

A secondary red line concerns asymmetry of expertise. If one party controls proprietary models or exclusive data, mediators must either restrict their use or provide compensatory support to other parties. Allowing unequal technical authority at the agenda stage distorts outcomes irreversibly.

Direct negotiation and bargaining

During direct negotiation, artificial intelligence has the narrowest legitimate role. Its use should be limited to internal support functions such as drafting assistance, comparative text analysis, and scenario visualization requested by all parties. AI should not independently generate proposals or evaluate concessions.

The governing design rule is human primacy. All proposals, tradeoffs, and formulations must be explicitly attributed to human judgment. AI outputs may inform drafts, but they must never be presented as recommendations in their own right.

The red line here is decisional delegation. Mediators must not allow AI systems to suggest optimal compromises, fair distributions, or equilibrium outcomes during bargaining. Negotiated settlement requires ambiguity, creativity, and political ownership. Algorithmic optimization undermines all three.

A second red line concerns performative authority. AI tools should not be displayed or invoked in ways that signal superior rationality or inevitability. The visual power of models must be carefully managed to avoid coercive framing.

Agreement drafting and validation

In drafting agreements, artificial intelligence offers speed and consistency. It supports clause comparison, legal pattern recognition, and internal coherence checks. These functions reduce error and accelerate convergence.

The design rule here is traceability. Every AI assisted draft element must remain fully traceable to human decision making. Mediators and parties must be able to explain why language appears as it does without reference to algorithmic authority.

The red line concerns authorship confusion. If parties cannot distinguish between politically agreed language and technically generated text, accountability erodes. Agreements must remain human artifacts, not system outputs.

Implementation, monitoring, and compliance

Post agreement, artificial intelligence becomes attractive for monitoring compliance, detecting violations, and tracking narratives that threaten implementation. This stage offers genuine benefits, especially in ceasefire monitoring and resource management regimes.

The design rule is proportionality. AI monitoring tools should match the sensitivity of the agreement and respect sovereignty, privacy, and consent. Monitoring should support confidence building, not enforcement by exposure.

The central red line is unilateral surveillance. Mediators must not deploy AI monitoring systems that one party perceives as externalized control or intelligence collection. Monitoring legitimacy depends on shared understanding and agreed scope.

A secondary red line concerns automatic escalation. AI detected violations must never trigger responses without human review. Automated inference must not replace political judgment in fragile post conflict environments.

Cross cutting rule. AI as political infrastructure, not political actor

Across all stages, one rule dominates. Artificial intelligence must function as political infrastructure rather than political authority. It shapes information environments, not outcomes. It informs judgment, not legitimacy.

The Ambio water diplomacy findings clarify why this matters. AI reshapes what counts as knowledge. In contexts of asymmetry, this reshaping redistributes power. Mediators therefore carry responsibility not only for how AI is used, but for how its authority is perceived.

The ultimate red line cuts across the lifecycle. AI must never substitute for trust. Once technical authority replaces relational legitimacy, mediation ceases to function as mediation.

Final synthesis

Mapped across the mediation lifecycle, artificial intelligence strengthens mediation when it expands participation, surfaces assumptions, and accelerates understanding while preserving contestation. It weakens mediation when it claims neutrality, concentrates expertise, or displaces political ownership.

International water diplomacy illustrates these dynamics in concentrated form. Resource negotiations amplify AI’s political effects early and visibly. Other mediation domains follow the same trajectory more gradually.

The implication for mediators is not rejection, but discipline. Artificial intelligence belongs in mediation only under conditions of transparency, symmetry, and human control. When these conditions fail, restraint becomes a professional obligation rather than a technical preference.

The use of AI in international diplomacy increasingly shapes power, norms, and bargaining leverage, not only administrative efficiency. Analysis published by the Konrad-Adenauer-Stiftung frames AI as a strategic factor in diplomacy, comparable to energy security or technological standards. States do not only compete over AI capabilities. They compete over who defines rules, ethics, and acceptable uses in global governance forums.

The article stresses that AI alters diplomatic practice in three concrete ways. First, it expands state capacity for analysis, forecasting, and narrative production. Diplomatic services already rely on automated translation, data aggregation, and pattern recognition to track crises, public opinion, and elite signaling. This accelerates decision cycles and rewards actors with stronger digital infrastructure. Second, AI becomes a subject of diplomacy itself. Questions of regulation, export controls, standards, and military or surveillance use now feature in bilateral and multilateral negotiations, especially between technologically advanced states and those seeking safeguards. Third, AI reshapes normative competition. Competing models of governance, liberal regulatory approaches versus state-centric or security-driven models, now play out through AI norms.

The Konrad-Adenauer-Stiftung analysis highlights a core diplomatic risk. AI systems project authority through technical credibility. This risks depoliticizing contested choices by presenting them as technical necessities. In diplomacy, this shifts debates from values and interests toward model outputs and expert claims. Smaller or less technologically advanced states risk exclusion if they cannot interrogate data sources, assumptions, or algorithmic design. This mirrors concerns seen in international water negotiations and resource diplomacy.

The article also emphasizes European priorities. From this perspective, AI governance in diplomacy must reflect transparency, accountability, and human oversight. Diplomatic legitimacy depends on explainable systems and clear responsibility chains. When AI informs policy positions or negotiation stances, diplomats remain accountable for outcomes. The technology does not absorb political responsibility.

Looking ahead, the Konrad-Adenauer-Stiftung argues that AI literacy will become a core diplomatic skill. Negotiators must question training data, funding sources, and embedded assumptions. Diplomatic services must invest in in-house expertise rather than outsourcing strategic judgment to private vendors or opaque systems. AI use in diplomacy will expand. Its political effects will expand faster.

The central implication aligns with broader conflict mediation research. AI strengthens diplomacy when it supports analysis, inclusion, and transparency. It undermines diplomacy

when it concentrates authority, obscures value choices, or replaces political judgment with technical certainty.

The use of artificial intelligence in international conflict mediation raises a harder question than simple efficiency. The core issue concerns power, judgment, and legitimacy. Mediation rests on trust, political consent, and negotiated interpretation of facts. AI reshapes each of these elements. The question therefore is not whether AI improves mediation in a technical sense. The question asks whether AI strengthens or weakens the political foundations that make mediation possible.

This analysis examines how AI is used today, which stages of mediation it touches, the advantages it offers, the risks it introduces, and whether its use makes sense on normative, practical, and substantive grounds. The assessment draws on current practice within multilateral diplomacy, peace support operations, and resource negotiations, including international water diplomacy.

How AI enters conflict mediation

AI enters mediation through the back door, not the negotiating table. Most deployments support analysis, preparation, and monitoring rather than direct bargaining. This pattern reflects a shared caution among mediators. Political judgment remains human. AI supports the informational environment around judgment.

Three institutions illustrate current practice: United Nations Department of Political and Peacebuilding Affairs, Stockholm International Peace Research Institute, and the Konrad-Adenauer-Stiftung. Each approaches AI as an analytical instrument embedded inside political processes rather than a decision-maker.

Within mediation support units, AI tools process large volumes of unstructured data. These include social media content, local news, satellite imagery, transcripts, and field reports. Analysts use machine learning techniques to identify narrative shifts, escalation signals, and actor networks. The goal focuses on speed and coverage rather than prediction certainty.

AI also supports translation and summarization across languages. This function matters in multi-party talks involving regional actors, armed groups, and external guarantors. Automated translation shortens feedback loops between field reporting and senior mediation teams.

A third use appears in scenario modeling. In climate-linked disputes such as water sharing, land use, or migration corridors, AI-driven models simulate future stress conditions. These models shape agenda setting by narrowing the range of plausible futures discussed during talks.

Areas of mediation affected by AI

AI affects mediation unevenly across stages. It exerts the strongest influence before talks begin and after agreements are signed.

Early warning and preventive diplomacy

Early warning systems represent the most mature application. AI scans patterns across time and geography. It flags correlations between rhetoric, mobilization, and violence. This function supports preventive diplomacy by prioritizing attention and resources.

Within UN political missions, analysts rely on AI-assisted tools to track narrative escalation. The system does not predict war. It highlights rising risk so diplomats intervene earlier. This intervention often takes the form of quiet engagement rather than formal mediation.

This use raises fewer legitimacy concerns because outputs remain internal. Parties do not negotiate over model results. Mediators retain full discretion.

Preparation and process design

AI shapes who sits at the table and which issues appear negotiable. Stakeholder mapping tools extract relational data from reports, social media, and prior agreements. These tools reveal informal brokers and peripheral actors.

This capacity offers value in fragmented conflicts involving militias, community leaders, and external sponsors. AI surfaces connections that traditional reporting misses. It also risks overemphasizing digitally visible actors at the expense of offline influence.

Agenda design increasingly reflects model outputs. Climate projections, demographic trends, or infrastructure simulations define problem boundaries. In international water negotiations, AI-driven hydrological models determine which tradeoffs enter discussion. This effect marks a political shift. Technical frames narrow political choice.

Negotiation support and drafting

Direct use of AI during talks remains limited. Some mediation teams use AI for rapid drafting of options papers, ceasefire language comparisons, or sequencing scenarios. Humans revise all outputs.

The value lies in speed and consistency. The risk lies in subtle framing effects. Draft language shapes expectations. Even neutral wording reflects assumptions embedded in training data and prior agreements.

Implementation and monitoring

Post-agreement monitoring represents a growing domain. AI analyzes compliance signals such as troop movements, ceasefire violations, or online incitement. These tools support verification mechanisms and confidence-building measures.

This use strengthens agreements by reducing reliance on self-reporting. It also raises sovereignty concerns when monitoring relies on external data sources or private vendors.

Examples from practice

International water diplomacy offers a clear example of AI's political impact. Research published in the journal Ambio documents how AI-driven hydrological modeling enters negotiations over shared river basins. More than 420 basins cross national borders. Power asymmetries define many of these negotiations.

AI models offer forecasting and visualization capabilities. States with limited technical capacity gain access to analytical tools through open platforms. This access improves bargaining confidence. At the same time, proprietary models funded by donors or upstream states introduce trust deficits. Negotiators struggle to contest outputs they do not understand or control.

Interviews with practitioners reveal concern over false neutrality. When one party presents AI forecasts as objective truth, political disagreement shifts into technical dispute. This shift favors actors with data access and modeling expertise.

Within UN preventive diplomacy, AI-assisted social media analysis already supports early intervention. Tools scan hate speech trends and mobilization rhetoric. Diplomatic teams receive structured briefings rather than raw data. This approach improves response time without displacing judgment.

In negotiation research, experimental work explores AI as a coaching instrument. These systems analyze negotiation behavior and suggest tactics. Practitioners treat them as training aids rather than operational tools. High-stakes mediation resists automation of interpersonal dynamics.

Advantages of AI in conflict mediation

Speed and scale

AI processes volumes of information that overwhelm human teams. Conflict environments generate constant data flows. AI reduces latency between signal and response. Faster synthesis supports timely diplomacy.

This advantage matters most in preventive contexts. Early engagement often prevents escalation. Delays cost leverage.

Analytical reach

AI expands analytical reach beyond elite reporting. It captures local narratives, diaspora discourse, and transnational mobilization. This breadth improves situational awareness.

For mediators working with small teams, AI substitutes for manpower. It lowers operational burden without expanding field presence.

Scenario exploration under uncertainty

Climate stress, demographic change, and economic shocks complicate mediation. AI-driven models explore conditional futures. These models support structured discussion of tradeoffs.

In water disputes, scenario modeling helps negotiators visualize drought or flood impacts across borders. This visualization supports problem-solving frames rather than zero-sum claims.

Language and access

Automated translation and summarization reduce linguistic barriers. This function improves inclusion of local actors and non-English sources. It also reduces reliance on intermediaries.

For multilateral processes, language support accelerates coordination and reduces misinterpretation.

Disadvantages and structural risks

Power asymmetry and exclusion

AI reflects existing power structures. Data availability skews toward connected regions and dominant languages. Actors with digital infrastructure gain disproportionate visibility.

In negotiations, this skew privileges certain narratives. Marginalized groups struggle to contest AI-informed claims. Exclusion deepens rather than narrows.

False objectivity

AI outputs project authority. Numbers and models imply neutrality. In political negotiations, this implication distorts debate. Parties argue over model validity instead of interests and values.

The water diplomacy evidence highlights this risk. Technical framing suppresses distributive conflict without resolving it.

Opacity and contestability

Many AI systems lack transparency. Proprietary models prevent scrutiny. Negotiators cannot interrogate assumptions or data sources.

This opacity undermines trust. Mediation depends on perceived fairness. Black-box tools erode legitimacy even when outputs align with reality.

Manipulation incentives

Once parties understand model inputs, they seek to game them. Coordinated information campaigns flood social media. Staged signals distort early warning systems.

AI raises the stakes of information warfare inside diplomacy. Mediators must treat outputs as contested terrain rather than neutral input.

Responsibility and accountability

AI complicates responsibility chains. When mediators rely on AI-informed analysis, attribution blurs. Political accountability remains human, yet decision pathways grow opaque.

This tension raises ethical concerns. Mistakes carry human cost. Mediators must justify decisions grounded partly in algorithmic processes.

Security and confidentiality

Mediation requires confidentiality. AI introduces data handling risks. Cloud services, vendors, and cross-border data flows expand exposure.

Sensitive negotiation materials demand strict control. Technical convenience conflicts with diplomatic prudence.

Normative assessment

Normatively, AI challenges core mediation principles. Mediation values consent, equality of voice, and political ownership. AI risks undermining each principle when introduced without safeguards.

Consent erodes when parties accept technical frames they cannot contest. Equality suffers when access and literacy diverge. Ownership weakens when models define outcomes.

At the same time, exclusion already defines many conflicts. AI offers tools to widen participation when deployed deliberately. Open systems and shared modeling processes support fairness.

The normative question therefore hinges on governance. AI does not violate mediation norms by default. Ungoverned AI does.

Effectiveness assessment

Effectiveness depends on function. AI improves effectiveness in analysis, preparation, and monitoring. Evidence from UN practice supports this claim.

Effectiveness declines when AI enters bargaining dynamics directly. Trust and legitimacy suffer. Political compromise requires ambiguity and flexibility. Models resist both.

AI works best as infrastructure rather than interface.

Substantive assessment

Substantively, AI reshapes what counts as knowledge. It privileges quantifiable phenomena. It sidelines historical grievances, identity, and emotion.

In resource conflicts, this shift proves especially consequential. Water flows, land use, and emissions lend themselves to modeling. Distributional justice does not.

Substantive outcomes improve when mediators recognize this limitation and supplement AI with qualitative engagement.

Conclusion

The use of AI in international conflict mediation makes sense under strict conditions. It strengthens mediation when it expands information access, accelerates synthesis, and supports early engagement. It weakens mediation when it concentrates authority, obscures political choice, or substitutes technical certainty for trust.

AI should remain a servant of mediation, not a source of legitimacy. Mediators must treat AI outputs as contested inputs, not neutral facts. They must invest in AI literacy, transparency, and shared access. They must retain responsibility for judgment.

International water diplomacy offers a warning and an opportunity. Where resources, data, and power intersect, AI amplifies politics rather than replacing it. Conflict mediation across domains follows the same logic.

Used carefully, AI supports peace. Used carelessly, it hardens inequality and mistrust. The difference lies in governance, not technology.

The use of artificial intelligence in international conflict mediation raises foundational questions about power, legitimacy, and political judgment. Mediation operates through consent, trust, and negotiated interpretation of contested realities. Artificial intelligence reshapes how information enters these processes and how authority attaches to knowledge. The central issue is not

whether artificial intelligence improves efficiency. The central issue concerns whether its use strengthens or weakens the political conditions under which mediation succeeds.

This essay analyzes how artificial intelligence is used in international conflict mediation, the specific stages of mediation it influences, the advantages it offers, the risks it introduces, and the normative, effective, and substantive implications of its use. The analysis draws on current practice in multilateral diplomacy, preventive diplomacy, post conflict implementation, and resource based negotiations, with particular attention to international water disputes.

Artificial intelligence enters conflict mediation indirectly rather than overtly. It rarely appears at the negotiating table. Its influence operates through preparation, analysis, monitoring, and agenda formation. This pattern reflects deliberate restraint by mediators and diplomatic institutions. Political authority remains human. Artificial intelligence functions as infrastructure that shapes how problems appear and which options seem plausible.

Within the United Nations Department of Political and Peacebuilding Affairs, artificial intelligence supports political analysis and early warning. Machine learning systems process large volumes of unstructured data, including social media content, local news reporting, satellite imagery, and field mission reports. Analysts use these systems to detect narrative shifts, escalation signals, and emerging actor networks. The purpose centers on prioritization rather than prediction. Artificial intelligence helps analysts decide where to focus attention and diplomatic effort.

Artificial intelligence also supports language processing and synthesis. Multilingual conflicts generate constant translation demands. Automated translation and summarization shorten feedback loops between field reporting and senior mediation teams. This capacity accelerates internal coordination and reduces reliance on small pools of human translators. It also shapes which information circulates upward within institutions.

Scenario modeling represents another entry point. In conflicts linked to climate stress, migration, land use, or shared natural resources, artificial intelligence driven models simulate future conditions under varying assumptions. These simulations inform briefing materials, options papers, and agenda setting. In international water disputes, hydrological modeling increasingly defines the factual baseline for negotiations.

The influence of artificial intelligence varies across the mediation lifecycle. Its strongest effects appear before formal talks begin and after agreements enter implementation.

Early warning and preventive diplomacy represent the most established domain. Artificial intelligence identifies correlations between rhetoric, mobilization, and violence across time and geography. Analysts receive structured outputs highlighting areas of rising risk. Diplomatic engagement often follows quietly through informal channels. This use raises fewer legitimacy concerns because outputs remain internal and advisory. Parties do not negotiate over model results. Mediators retain discretion over interpretation and response.

During preparation and process design, artificial intelligence shapes participation and agenda structure. Stakeholder mapping tools extract relational data from reports, online discourse, and

prior agreements. These tools surface informal brokers and peripheral actors who might escape traditional reporting. In fragmented conflicts involving militias, local authorities, and external sponsors, this mapping expands situational awareness. At the same time, digital visibility skews representation. Actors with limited online presence risk marginalization.

Agenda formation increasingly reflects model outputs. Climate projections, demographic trends, and infrastructure simulations define problem boundaries. In international water negotiations, artificial intelligence driven hydrological models determine which tradeoffs enter discussion. These models influence perceptions of scarcity, risk, and responsibility. Technical framing narrows political choice by constraining which futures appear negotiable.

Direct use of artificial intelligence during bargaining remains limited. Some mediation teams employ artificial intelligence to assist rapid drafting of options papers, ceasefire clauses, or sequencing scenarios. Human mediators revise all outputs. The value lies in speed and consistency. The risk lies in framing effects. Draft language shapes expectations and signals priorities. Training data and prior agreements embed assumptions that influence these drafts.

Post agreement implementation and monitoring represent a growing area of use. Artificial intelligence analyzes compliance indicators such as troop movements, ceasefire violations, or online incitement. These tools support verification mechanisms and confidence building measures. Monitoring based on external data sources strengthens enforcement while raising sovereignty and privacy concerns.

Concrete examples illustrate these dynamics. International water diplomacy offers a clear case of artificial intelligence shaping political outcomes. Research conducted by scholars affiliated with the Stockholm International Peace Research Institute documents how artificial intelligence driven hydrological modeling enters negotiations over shared river basins. More than four hundred river basins cross national borders. Many involve pronounced power asymmetries between upstream and downstream states.

Artificial intelligence based models provide forecasting and visualization capabilities that influence bargaining positions. Open access tools enable states with limited technical institutions to engage with complex data. This access improves confidence and participation. At the same time, proprietary models funded by donors or dominant states introduce trust deficits. Negotiators struggle to contest outputs they neither control nor fully understand.

Interviews with practitioners reveal concern over false neutrality. When one party presents artificial intelligence forecasts as objective truth, political disagreement shifts toward technical dispute. This shift advantages actors with modeling expertise and data access. Underlying distributive conflicts persist without resolution.

Within United Nations preventive diplomacy, artificial intelligence assisted analysis already supports early engagement. Systems scan online discourse for hate speech trends and mobilization rhetoric. Diplomatic teams receive synthesized briefings rather than raw data. This approach improves response time while preserving human judgment.

In academic and training contexts, artificial intelligence appears as a negotiation support and coaching tool. Experimental systems analyze negotiation behavior and suggest tactics. Practitioners treat these systems as educational aids rather than operational instruments. High stakes mediation resists automation of interpersonal dynamics, trust building, and moral authority.

Artificial intelligence offers several advantages in conflict mediation. The most significant advantage concerns speed and scale. Conflict environments generate vast and continuous data flows. Artificial intelligence processes this volume faster than human teams. Reduced latency between signal and response supports timely diplomacy. Preventive engagement benefits most from this acceleration.

Analytical reach expands through artificial intelligence. Systems capture local narratives, diaspora discourse, and transnational mobilization that escape elite reporting. This breadth improves situational awareness. For small mediation teams, artificial intelligence substitutes for manpower without expanding field presence.

Scenario exploration under uncertainty represents another advantage. Climate stress, demographic change, and economic shocks complicate negotiation. Artificial intelligence driven models explore conditional futures and tradeoffs. In water disputes, scenario visualization supports problem solving approaches rather than zero sum framing.

Language processing improves access and inclusion. Automated translation and summarization reduce linguistic barriers. Local actors gain greater visibility. Multilateral coordination accelerates. Misinterpretation decreases.

These advantages depend on conditions. Artificial intelligence delivers value when access remains broad, assumptions remain visible, and human judgment remains central.

The disadvantages and risks of artificial intelligence in conflict mediation carry structural significance. Power asymmetry presents the most serious risk. Artificial intelligence reflects existing inequalities in data availability, language dominance, and technical capacity. Digitally connected actors gain disproportionate visibility. Marginalized groups struggle to contest artificial intelligence informed claims. Exclusion deepens rather than narrows.

False objectivity compounds this risk. Artificial intelligence outputs project authority through numbers, models, and visualizations. In political negotiations, this projection distorts debate. Parties argue over model validity rather than interests and values. Technical framing suppresses distributive conflict without resolving it.

Opacity undermines contestability. Proprietary systems prevent scrutiny. Negotiators lack access to training data, funding sources, and model design. Trust erodes when claims rest on black box processes. Mediation depends on perceived fairness. Opaque tools weaken legitimacy even when outputs align with reality.

Manipulation incentives emerge once parties understand model inputs. Coordinated information campaigns flood social media channels. Artificial signals distort early warning systems. Artificial intelligence becomes another terrain for information warfare. Mediators must treat outputs as contested rather than neutral.

Responsibility and accountability grow harder to trace. Artificial intelligence complicates decision chains. Political responsibility remains human, yet analytical pathways grow opaque. Ethical concerns follow. Mistakes carry human cost. Mediators must justify decisions grounded partly in algorithmic processes.

Security and confidentiality present additional challenges. Mediation relies on discretion. Artificial intelligence introduces data handling risks through vendors, cloud infrastructure, and cross border data flows. Sensitive negotiation materials demand strict control. Technical convenience conflicts with diplomatic prudence.

Normatively, artificial intelligence challenges core principles of mediation. Mediation values consent, equality of voice, and political ownership. Artificial intelligence undermines these principles when introduced without safeguards. Consent weakens when parties accept technical frames they cannot contest. Equality suffers when access and literacy diverge. Ownership erodes when models define outcomes.

At the same time, exclusion already defines many conflicts. Artificial intelligence offers tools to widen participation when deployed deliberately. Open systems, shared modeling processes, and capacity building support fairness. The normative assessment therefore depends on governance rather than technology itself.

Effectiveness varies by function. Artificial intelligence improves effectiveness in analysis, preparation, and monitoring. Evidence from United Nations practice supports this conclusion. Effectiveness declines when artificial intelligence enters bargaining dynamics directly. Trust and legitimacy suffer. Political compromise relies on ambiguity and flexibility. Models resist both. Artificial intelligence works best as infrastructure rather than interface.

Substantively, artificial intelligence reshapes what counts as knowledge. It privileges quantifiable phenomena and measurable variables. Historical grievance, identity, and emotion receive less weight. In resource conflicts, this shift proves consequential. Water flows, land use, and emissions lend themselves to modeling. Justice and historical responsibility do not. Outcomes improve when mediators supplement artificial intelligence with qualitative engagement and political dialogue.

A broader diplomatic perspective reinforces these conclusions. Analysis by the Konrad-Adenauer-Stiftung frames artificial intelligence as a strategic factor in diplomacy rather than a neutral tool. States compete over artificial intelligence governance, standards, and legitimacy. Artificial intelligence reshapes not only mediation practice but the normative environment in which mediation occurs. Smaller states risk exclusion when technical authority replaces political argument.

The final assessment must therefore remain conditional. The use of artificial intelligence in international conflict mediation makes sense under strict constraints. It strengthens mediation when it expands information access, accelerates synthesis, and supports early engagement. It weakens mediation when it concentrates authority, obscures political choice, or substitutes technical certainty for trust.

Artificial intelligence should remain subordinate to mediation rather than a source of legitimacy. Mediators must treat artificial intelligence outputs as contested inputs rather than neutral facts. Investment in artificial intelligence literacy, transparency, and shared access remains essential. Responsibility for judgment must remain explicit and human.

International water diplomacy illustrates both warning and opportunity. Where resources, data, and power intersect, artificial intelligence amplifies politics rather than replacing it. Conflict mediation across domains follows the same logic. Artificial intelligence supports peace when governance restrains its authority. Absent such restraint, it entrenches inequality and mistrust. The difference lies in political design rather than technological capability.