AI Negotiation Forum 2026

10 – 11 January 2026

Exploring Pathways for Responsible Use of AI
in Negotiation and Diplomacy

<u>Final Report</u>

A collaboration between

Table of Contents

AI Negotiation Forum 2026
10 – 11 January 2026

Exploring Pathways for Responsible Use of AI
in Negotiation and Diplomacy

Executive Summary

Artificial intelligence is rapidly transforming how negotiators, mediators, and diplomats prepare, analyse, and engage in complex negotiations. Yet this transformation is unfolding faster than reflection, training, and governance frameworks can keep pace.

The AI Negotiation Forum was convened to address this gap. Anchored in the lived experience of practitioners, the Forum brought together negotiators, researchers, technologists, and educators to examine how AI is actually being used in negotiation and diplomacy, what value it creates, what risks it introduces, and what capacities and safeguards are urgently needed.

This report synthesises the findings of the Forum's first edition, drawing on a global survey of practitioners, focused qualitative discussions, and expert roundtables. It captures observed practices, practitioner reflections, and expert analysis across four core areas: research and development, capacity building, risk management, and the impact of AI on negotiation, mediation, and diplomacy.

Key messages from the Forum

- AI is already widely used by negotiators, primarily for preparation, analysis, drafting, and sense-making, but adoption is largely informal and weakly governed.

- Practitioners report clear gains in efficiency and analytical support, while remaining cautious about bias, confidentiality, over-reliance, and loss of human judgment.

- The main risks of AI in negotiation are cognitive, ethical, and institutional rather than purely technical, and are amplified by gaps in training and governance.

- Human judgment, accountability, and relational skills remain central and cannot be delegated to AI systems; future tools must support human–AI collaboration rather than automation.

- There is an urgent need for coordinated action across practitioners, institutions, researchers, and training providers to build capacity, manage risks, and guide responsible innovation.

By grounding its analysis in practice rather than conceptual abstraction, this report aims to inform and guide responsible innovation, strengthen human judgment, and support institutions and practitioners in shaping AI as a tool that serves, rather than reshapes or replaces, the human foundations of negotiation.

* * *

Foreword

The **AI Negotiation Forum** is the result of a sustained and collaborative effort bringing together the **Frontline Associates Training Center (FATC)** and a network of internationally recognised academic and policy institutions. Conceived as a practice-driven initiative, the Forum aims to anchor reflection on artificial intelligence in the lived experience of negotiators, mediators, and diplomats, while maintaining a high level of analytical and scientific rigor.

The authors of this report wish to express their deep gratitude to the members of the Forum's **Governing Committee**, in particular:

- **Ma Zilin, PhD**, Postdoctoral Researcher, Digital Data Design (D^3) Institute at Harvard
- **Delphine Alles, PhD**, Professor of Political Sciences, Vice-President, INALCO; Scientific Director, DÉCRIPT
- **Francesco Marchi, PhD**, Director of Negotiation Expertise at ALTERNEGO; Lecturer, Sciences Po (Paris), College of Europe (Bruges), and University Paris 1 Panthéon-Sorbonne
- **Annie Chaloux, PhD**, Associate Professor, Université de Sherbrooke; Holder of the Quebec Research Chair in Climate Scientific Diplomacy

Since the summer of 2025, they have worked assiduously to guide the Forum's design, methodology, and intellectual direction. Their sustained engagement ensured that the deliberations remained both closely connected to practice and firmly grounded in scholarly standards.

The authors would also like to thank warmly **Martin Wählisch**, Associate Professor at the University of Birmingham, for his support and guidance in the development of the survey questionnaire, as well as all the organisations that contributed to the dissemination of the survey within their respective networks. Sincere appreciation is extended to the contributors to the roundtables and to the presenters at the Innovation Exchange Lab, whose generous engagement helped shape and enrich the Forum's discussions.

Particular appreciation is extended to the **Harvard Kennedy School of Government** for hosting the Forum as part of the annual *Frontline Negotiation Lab* course. This institutional anchoring provided a unique environment in which practitioners, students, and researchers could engage in structured reflection and dialogue at the intersection of negotiation practice, public policy, and emerging technologies.

The Forum also benefited greatly from the commitment and professionalism of **Harvard graduate students and mid-career fellows**, whose contributions were essential to its success. Their work in organising the Forum, facilitating focus group

discussions, and supporting the roundtables enabled rich, disciplined exchanges. Their participation, insights, and comments significantly informed and sharpened the analysis presented in this report.

Finally, I wish to extend special thanks to the **FATC team**, **Suzi Jazmati**, Community Manager; **Dariha Erketaeva**, Negotiation Specialist; **Charles McJilton**, Senior Advisor; **Eloïse Roy**, Partnership Coordinator; and **Mathilde Germain**, Researcher, who are at the core of FATC's vision and mission and whose sustained efforts made this initiative possible.

This report reflects a shared endeavour. It is offered as a contribution to ongoing dialogue and collective learning on how artificial intelligence can be responsibly integrated into negotiation, mediation, and diplomacy, strengthening, rather than undermining, human judgment, legitimacy, and accountability.

Claude Bruderlein
Director, Frontline Associates Training Center

Objectives of the AI Negotiation Forum

The AI Negotiation Forum is guided by a set of interrelated objectives that reflect both the realities of current practice and the strategic challenges posed by the rapid integration of artificial intelligence into negotiation and diplomacy.

These objectives define the orientation of the AI Negotiation Forum toward a broader and growing community of practitioners and researchers engaged with the use of artificial intelligence in negotiation, mediation, and diplomacy. They are intended not only to guide the Forum's own activities, but also to provide a shared reference point for institutions, training programs, research initiatives, and professional networks seeking to engage in this emerging field.

By articulating these objectives openly, the Forum positions itself as a convening platform for collective learning, experimentation, and standard-setting, inviting practitioners and researchers to contribute to, challenge, and further develop responsible approaches to AI-assisted negotiation.

1. *Bridge research, theory, and field practice*

Connect AI research and negotiation theory with frontline experience to ensure that technological development is grounded in operational reality, responds to real constraints, and addresses practitioner needs rather than abstract optimization goals.

2. *Identify and mitigate risks at the human–AI interface*

Analyse and address risks related to AI use in negotiation, including bias, misuse, over-reliance, data sensitivity, and accountability gaps. Develop practical safeguards, ethical guardrails, and operational standards to manage these risks.

3. *Contribute to the development of shared standards and norms*

Contribute to the definition of common principles, ethical standards, and good practices for AI-assisted negotiation and mediation, informed by cross-sector experience and tested in real-world settings.

4. *Foster cross-sector collaboration and dialogue*

Create a structured space for academics, practitioners, technologists, and policymakers to jointly question assumptions, share evidence, and co-design AI approaches that are credible, usable, and aligned with public interest goals.

5. *Enhance AI's contribution to major global challenges*

Direct AI innovation and practice toward challenges where negotiation plays a central role, such as conflict resolution, humanitarian access, diplomacy, and collective action, seeking measurable improvements in outcomes and impact.

Forum Methodology

The Forum was designed as a step-by-step pathway to bridge AI negotiation practices to AI theory and research. It started from how negotiators actually use AI in practice and progressively translated this experience into research priorities, capacity building agendas, risk management, and overarching impact on negotiation and diplomacy.

Gathering data on AI negotiation practices

The first step of the Forum process consisted of capturing current practice by experienced negotiators. A structured survey documented how negotiators use AI in their daily work, the benefits they identified, the limitations they encountered, and their concerns regarding risks, ethics, and integration into professional workflows. The AINF Survey was developed in partnership with established academic institutions and conducted under the supervision of a scientific committee to ensure methodological rigor, relevance, and independence. The survey was circulated between 19 November and 20 December, collecting 485 responses, primarily from mid- to senior-level negotiators, ensuring a high level of professional relevance (see the analysis of the survey results in Annex).

The scientific committee overseeing the survey brought together academic, professional, and practitioner expertise from the Harvard School of Engineering and Applied Sciences, the Geneva Graduate Institute, INALCO, the European Negotiation Association, the University of Sherbrooke, the University of Birmingham, and the FATC team (see Annex X). This governance structure ensured that the evidence base reflected both scholarly standards and field realities.

Reflecting on the challenges of the use of AI in negotiation

The second step involved analysing and deepening reflection on this practice. Focus groups composed of selected respondents were held on 10 January 2026, hosted by the Harvard Kennedy School of Government and facilitated by mid-career students. More than 50 survey respondents from diverse background and regions participated. These discussions enabled practitioners to reflect collectively on how AI affects judgment, workflow, trust-building, and professional identity, and to articulate unmet needs and concerns.

Practitioner reflections were structured around four agenda areas, which directly informed the roundtables held on 11 January 2026:

- Research and development needs, including methods, tools, and strategies for AI-supported negotiation;
- Capacity building for negotiators to use AI in a safe, effective, and responsible manner;
- Identification and mitigation of cognitive, ethical, and institutional risks; and
- Ways to enhance the positive impact of AI in diplomacy and negotiation.

Deliberating on the future of AI in negotiation and diplomacy

The fourth step brought practitioner perspectives into dialogue with experts. The roundtables took place on 11 January 2026, hosted by the Harvard Kennedy School of Government. They were moderated by members of the Forum's scientific committee, ensuring continuity between the evidence gathered, practitioner reflection, and expert discussion.

More than twenty contributors, including experts, scientists, academics, and senior practitioners, shared their perspectives across the four agenda areas identified by the focus groups. To anchor discussions in practitioner experience, each contributor received preliminary results from the survey as well as a synthesised summary of the focus group discussions in advance. These materials informed their interventions and ensured that expert perspectives responded directly to observed practices, needs, and concerns.

The roundtables convened a broad and diverse audience, with over 250 professionals, experts, and graduate students participating online in the Forum on 11 January. In addition to the roundtable discussions, the Forum included a series of exhibits and demonstrations, showcasing emerging startups, new AI-supported negotiation platforms, and applied tools. These exhibits highlighted AI use across multiple negotiation domains, including climate negotiations, humanitarian negotiations, commercial negotiations, and political negotiations. Together, the roundtables and exhibits reinforced the Forum's objective of connecting practitioner experience, research insight, and technological innovation within a shared, practice-driven learning space.

Outline of the Report

Section I: Survey Findings and Practitioner Profile

The report opens with a presentation of the main observations emerging from the survey. This section synthetises key patterns in the use of AI in negotiation, perceived benefits, limitations, and concerns, and provides an overview of the professional profile of respondents, with particular attention to experience level, roles, and negotiation contexts.

Section II: Practitioner Reflections from the Focus Groups

The second section presents the main reflections expressed by participants in the focus groups. It deepens and contextualizes the survey findings by highlighting how practitioners interpret their own use of AI, articulate needs and risks, and frame expectations. These reflections served as the anchor for the deliberations conducted during the roundtables.

Section III: Roundtable Deliberations

The third section summarizes the discussions held in each roundtable. It captures how experts, researchers, and practitioners engaged with practitioner-driven insights across the four agenda areas identified by the focus groups, and how these discussions translated experience into research, training, and governance considerations.

Section IV. Conclusions and Next Steps

The final section draws together the key cross-cutting insights from the survey, focus groups, and roundtables. It outlines overarching conclusions and proposes next steps for research, capacity building, risk management, and future editions of the AI Negotiation Forum.

* * *

Section I: Survey Findings

Profile, Use of AI, and Practitioner Observations
(See the complete analysis of the survey results in the Annex)

Profile of Respondents

The AI Negotiation Forum Survey collected responses from 485 practitioners, primarily mid- to senior-level negotiators, mediators, and diplomats, operating in complex and high-stakes environments. Respondents are predominantly professionals with substantial experience in negotiation, with most reporting more than eight years of practice, giving the dataset a high level of professional depth and credibility.

Participants represent a broad range of roles, including senior managers, field and programme officers, professional mediators, consultants, policy advisors, and analysts. This diversity reflects a practitioner community that combines operational engagement with strategic responsibility. Respondents are affiliated mainly with international and regional organisations and international NGOs, complemented by representation from academic institutions, civil society organisations, private companies, and government agencies.

Geographically, the survey reflects the realities of negotiation-intensive contexts. Africa and the Middle East are the most represented regions, together accounting for more than half of responses, followed by Europe, largely through institutional, diplomatic, and research roles. Asia and the Americas appear as secondary but significant clusters. A large proportion of respondents work in global or multi-regional roles, consistent with multilateral, humanitarian, and diplomatic negotiation environments.

Use of AI in Negotiation Practice

Survey results indicate that AI is already widely integrated into professional workflows. Approximately two-third of respondents use AI daily or several times per week, while the majority report regular use across multiple tasks. Familiarity with AI tools increases steadily with use, suggesting rapid diffusion of applied AI practices among negotiators.

Learning how to use AI remains largely self-directed. Most respondents report acquiring skills through experimentation, online resources, peer exchanges, or task-specific problem solving rather than through structured institutional training. This pattern highlights a mismatch between individual adoption and institutional support.

In terms of tools, ChatGPT is the most widely used platform, followed by Microsoft Copilot, Google Gemini, and smaller clusters using Claude and Perplexity. Use of other tools remains marginal. This concentration suggests that negotiators rely primarily on general-purpose, widely accessible AI systems rather than bespoke negotiation technologies.

AI is most frequently used for summarising information, drafting documents and briefs, analysing contexts and risks, and exploring strategic options and scenarios. These uses indicate that AI is primarily valued for sense-making, preparation, and communication rather than autonomous decision-making.

Observations on the Usefulness of AI

Respondents report strong perceived benefits, particularly in terms of efficiency and cognitive support. Nearly 80 percent describe a concrete moment in which AI use led to a noticeable shift in their professional practice.

Key areas of usefulness include:

- Efficiency and speed, allowing practitioners to meet tight deadlines and reduce time spent on labour-intensive tasks;

- Analytical depth and decision support, where AI helps structure arguments, identify patterns, and broaden strategic perspectives;

- Writing and communication support, with AI often used as a starting point for reports, briefs, and talking points;

- Learning and upskilling, using AI as a tutor to acquire new competencies more quickly.

At the same time, respondents consistently emphasize the need for verification and human judgment. High agreement levels are reported for double-checking AI outputs, combining AI input with human intuition, and maintaining personal responsibility for AI-assisted outcomes. This reflects a strong culture of critical use rather than delegation.

Concerns and Perceived Risks

Concerns about AI are widespread and multifaceted. The most strongly expressed risks relate to:

- Confidentiality and data sensitivity;
- Bias, misinterpretation, and hallucinations;
- Loss of creativity, reflection, or independent thinking.
- Additional qualitative responses reinforce four categories of risk:
- Cognitive risks, including overreliance and de-skilling;
- Ethical risks, such as bias, misinformation, and lack of accountability;
- Systemic risks, including power asymmetries, opacity, and dependency;
- Operational risks, such as missing context or poor judgment in complex environments.

A large majority of respondents report having encountered inaccurate or misleading AI outputs, reinforcing the need for safeguards and critical oversight.

Trust in AI remains pragmatic and conditional. Respondents are willing to adjust their thinking when AI offers useful perspectives, but trust develops slowly and is grounded in perceived utility and the ability to override AI, rather than in confidence in the technology itself.

Needs and Expectations

A consistent finding is the gap between individual use of AI and institutional readiness. Respondents rate organisational support, training, leadership guidance, peer learning, and workflow integration, significantly lower than their own level of AI use. This gap is perceived as a source of risk and vulnerability.

Training and technical support emerge as the dominant needs. Respondents prioritise:

- Practical skills for effective AI use and workflow integration;
- Guidance on managing risks related to confidentiality, bias, and misinformation;
- Ethical frameworks and safeguards appropriate to sensitive negotiation contexts.

Foundational technical literacy is considered important but secondary to applied skills. Very few respondents report needing no training, indicating broad recognition that AI use in negotiation requires structured support.

Finally, respondents express cautious ambivalence about the formal role of AI in multilateral diplomacy. While some acknowledge potential opportunities, most highlight risks related to legitimacy, transparency, political bias, equity of access, and loss of human judgment. This reinforces the need for prudence, governance, and capacity building before deeper institutional integration of AI occurs.

Section II: Practitioner Reflections from the Focus Groups

This section presents the main reflections emerging from the focus group discussions held on 10 January 2026. Drawing on practitioner experience, the discussions deepened and interpreted survey findings and served as the primary anchor for the deliberations at the roundtables. Reflections are organised around the four agenda areas identified collectively by participants.

1. Research and Development: Making AI Safe and Effective for Negotiation Practice

Participants emphasized that AI is already embedded in negotiation practice, but primarily as a preparatory and support tool rather than as an instrument used during live negotiations. Current uses focus on background research, document analysis, information structuring, and stress-testing arguments. Use of AI in real-time negotiation settings remains limited, largely due to concerns about contextual sensitivity, reliability, and trust.

A central reflection was that existing AI systems are poorly suited to the human, relational, and political dimensions of negotiation. Participants repeatedly stressed that AI does not adequately capture emotions, trust-building, informal power dynamics, or tacit knowledge developed through experience. As a result, **practitioners strongly rejected the idea of AI replacing human judgment and instead framed future research priorities around human–AI collaboration.**

From a research and development perspective, participants called for AI systems that function as decision-support or "heads-up display" tools, designed to preserve productive friction rather than optimize away deliberation. Key priorities included improving transparency around data sources and model behaviour, reducing bias and misinformation, and designing systems that explicitly support reflection, scenario exploration, and internal challenge. Evaluation criteria for AI in negotiation, participants argued, should extend beyond accuracy and speed to include impacts on judgment, responsibility, and professional identity.

2. Capacity Building: Enabling Safe, Effective, and Responsible Use of AI

Discussions on capacity building revealed that most practitioners learn to use AI in informal, improvised, and necessity-driven ways, often under acute operational pressure. **AI skills are developed through experimentation and workarounds rather than structured training, resulting in uneven competence and limited strategic mastery.**

Participants identified a major gap between existing AI training and the realities of negotiation practice. Most training assumes stable institutions, clean data, and formal authority structures, while practitioners operate in messy, informal, and rapidly evolving environments. A critical missing competence is the ability to translate complex field realities into problem formulations that AI systems can meaningfully process.

The group strongly cautioned against an overemphasis on prompt engineering. Instead, **participants highlighted the need for judgment-centred AI literacy, including knowing when not to use AI, how to interpret and contest outputs, and how to maintain ethical and professional accountability**. Data governance and confidentiality emerged as central capacity-building issues, as fear of data exposure limits the use of more advanced AI applications.

Equity concerns were also prominent. Participants warned that unequal access to models, infrastructure constraints, and underrepresentation of Global South contexts risk reinforcing existing power imbalances. **Capacity building was therefore framed not only as a technical issue, but also as a matter of inclusion, representation, and professional agency.**

3. Risk Identification, Prevention, and Management

The risk-focused discussions underscored that AI-related risks are already materializing in practice. Participants framed these risks less as technical malfunctions and more as cognitive, ethical, and institutional vulnerabilities arising from how AI is trusted, used, and governed.

A dominant concern was misplaced trust in opaque systems. The black-box nature of AI models, combined with limited understanding of data provenance and system behaviour, creates discomfort in sensitive negotiation environments, particularly in humanitarian and conflict settings where errors carry significant consequences. Bias and contextual distortion were identified as structural risks, especially in regions and languages that are underrepresented in training data.

Over-reliance and automation bias emerged as acute dangers, particularly under time pressure. Participants described situations in which AI-generated outputs appeared coherent and confident but were strategically flawed or outdated. In fast-changing environments, AI's reliance on static or lagging data creates a false sense of situational awareness.

Institutional gaps amplify these risks. **Participants highlighted the absence of shared standards, training, and review mechanisms, which shifts responsibility and liability onto individual users.** While some partial mitigation strategies exist, such as closed systems or retrieval-augmented models, these remain fragmented and unevenly applied.

4. Enhancing the Impact of AI in Negotiation, Mediation, and Diplomacy

Discussions on impact focused on how AI is reshaping power, meaning, and professional roles in negotiation and diplomacy. Participants agreed that AI already influences how problems are framed, arguments structured, and realities interpreted, often without explicit acknowledgment or shared norms.

AI was seen as simultaneously democratizing information access and reinforcing structural asymmetries. While it can reduce informational disadvantage in real time, unequal access to advanced tools and secure infrastructure creates new forms of power imbalance. Participants from humanitarian and Global South contexts stressed that access to AI itself is becoming an object of negotiation.

Concerns were raised about AI's impact on trust and meaning making. AI optimizes for plausibility and efficiency, not for relational depth or ethical nuance. Overreliance risks flattening complexity and undermining processes where patience, legitimacy, and human presence are essential. Participants strongly rejected granting AI a formal or autonomous role in negotiations, instead advocating for transparency about AI use and agreement on safeguards as part of the negotiation process itself.

Across discussions, a shared view emerged: AI should function as an instrument or co-pilot, not an actor. Human judgment, accountability, and moral agency must remain central. Enhancing AI's impact in diplomacy therefore depends not on expanding autonomy, but on embedding AI use within ethical frameworks, professional norms, and negotiated rules of engagement.

Section III: Roundtable Deliberations: From Practitioner Insight to Collective Analysis

This section presents a synthesis of the deliberations held during the four roundtables convened on 11 January 2026. Building directly on the survey findings and the practitioner reflections articulated during the focus groups, the roundtables brought together negotiators, researchers, technologists, and policy experts to engage with practitioner-driven concerns and expectations.

Each roundtable was structured around one of the four agenda areas identified by the focus groups: research and development, capacity building, risk identification and management, and enhancing the impact of AI in negotiation, mediation, and diplomacy. Contributors were provided in advance with preliminary survey results and synthesised focus group reflections, ensuring that expert analysis responded directly to observed practice rather than abstract or purely technical considerations.

The discussions translated practitioner experience into forward-looking perspectives on research priorities, training models, governance frameworks, and institutional responsibilities. Rather than seeking consensus or definitive solutions, the roundtables aimed to surface key tensions, design principles, and open questions that will shape future work at the intersection of AI and negotiation.

The subsections that follow present a structured summary of each roundtable, capturing the main lines of analysis, points of convergence and divergence, and the implications identified by participants for research, practice, and policy.

Advancing next-generation systems for safe and reliable AI-supported negotiation

Focus of the Roundtable:

Where does AI meaningfully augment complex negotiation practice today, where does it fail structurally, and what does this imply for next-generation system design?

Moderator: **Zilin Ma**, Postdoctoral Researcher, Digital Data Design (D^3) Institute at Harvard

Contributors and themes:

Usmaan Ahmad CEO & Co-Founder, Expeditionary	*AI Architecture for Complex Negotiations: From Individual Augmentation to Collective Intelligence*
Yadvinder S. Rana Professor of Cross-Cultural Negotiation; Founder, NegoAI	*Prescriptive Agent Scaffolding for Reliable AI Negotiation Support*
Przemysław Kępczyński PhD Candidate, Institute of Law Studies, Polish Academy of Sciences **Aneta Napieralska** Attorney and Mediator; Member, Institute for Global Negotiation (Zurich)	*Mediator-Centric Model AI Architecture for Diplomatic Negotiations*
Hong Zhang Postdoctoral Researcher, Leuphana University Lüneburg	*AI in the Age of Complex Negotiation: A Human–AI Integrated Perspective*
Dr. Cornelia C. Walther Senior Fellow; Founder, Hi@Hub Harvard Learning and Innovation Lab. Wharton Neuroscience Initiative. Sunway Sunway University Center for Planetary Health	*ProSocial AI: What It Is and Why It Matters*

Background

Negotiators are increasingly relying on artificial intelligence for their preparation, analysis, and scenario development. However, institutional adoption remains limited and uneven, reflecting persistent concerns around the lack of trust towards AI models,

limited literacy, reliability, and proper governance. Complex negotiations, particularly those involving multiple actors and political or strategic ambiguity, reveal structural limitations in current AI systems. This roundtable examined whether these limitations stem from immature technology or from deeper architectural mismatches between how AI systems are designed and how negotiation actually functions in practice.

Core themes emerging

A strong convergence emerged among contributors around the view that AI should function as a collaborative and augmentative partner rather than an autonomous decision-maker. Participants emphasized that negotiation remains a fundamentally human activity, requiring judgment, interpretation, and responsibility that cannot be delegated to machines.

Several contributors highlighted a fundamental architectural mismatch between current large language model–based systems and the realities of diplomatic and complex negotiations. While LLMs are optimized for semantic coherence and consistency, negotiations often rely on strategic ambiguity that deliberately sustain incoherence. Applying AI-led coherent, unified narratives to situations that require signalling, bluffing, and implicit positioning can distort rather than support complex negotiation dynamics. This mismatch produces fragility and limits trust in AI-generated outputs. Several contributors implicitly described a hybrid intelligence model, in which human judgment and machine analysis operate as complementary capacities, each compensating for the other's structural limitations rather than competing for control.

The issue of reliability featured prominently throughout the discussion. Hallucinations, inconsistent reasoning, and memory loss were not treated as temporary defects but as structural deficiencies of existing models. Participants argued that efforts to improve reliability must therefore be engineered around AI systems *through external scaffolding*, including verification mechanisms and constrained workflows, rather than assumed to emerge through model improvement alone.

Another recurring theme concerned the institutional nature of negotiation. Negotiations are conducted by teams operating across strategic, tactical, and temporal levels, yet most AI systems remain designed for individual use. The absence of institutional memory and multi-user coordination capabilities was identified as a major design gap.

Finally, contributors emphasized the asymmetry between human and machine strengths. AI systems excel at processing large volumes of information, mapping options, and identifying patterns under complexity. Humans, by contrast, retain their superiority in contextual sense-making, relational judgment and ethical reasoning, capabilities that are central to effective negotiation processes.

Key tensions

The discussion identified a persistent tension between efficiency and agency. While AI can boost productivity by significantly accelerating analysis and facilitating preparation, participants warned that excessive reliance on AI models risks cognitive offloading and long-term erosion of professional judgment.

Participants also noted a widening gap between rapid technological innovation and slower institutional adoption. Governance frameworks, training models, and organizational readiness are lagging behind technical capabilities, creating uneven and sometimes risky patterns of use.

This tension highlights the broader challenge of maintaining agency amid AI, ensuring that negotiators remain active sense-makers and accountable actors rather than passive recipients of machine-generated reasoning. Preserving agency thus becomes not a secondary concern, but a core performance requirement of negotiation-support systems.

Finally, the limits of data-driven systems were highlighted. Much of what matters in negotiation, tacit knowledge, informal exchanges, political sensitivities, remains non-digital or deliberately undisclosed, constraining what AI systems can reliably capture or infer.

Operational and policy implications

The roundtable underscored the need to design AI systems around robust human-led processes, ensuring that minority views, points of friction, informal deliberation, and accountability issues are preserved rather than eliminated. Reliability should be treated as an architectural and workflow challenge, addressed through layered verification, memory structures, and rule-based constraints.

At the institutional level, effective deployment requires proper governance frameworks that address data protection, explainability, trust, and accountability. Capacity-building efforts should move beyond basic tool use toward deeper AI literacy for negotiators, focusing on limitations, failure modes, and ethical risks alongside potential benefits.

Open questions / areas for further inquiry

Key questions remain regarding how AI systems can maintain contextual and institutional memory across long negotiation processes without compromising confidentiality. Further research is needed to identify architectural models capable of supporting multi-party, multi-level negotiations involving competing domestic and international constraints.

Participants also questioned how success in AI-supported negotiation should be defined beyond speed and efficiency, and what forms of professional training are required to prevent de-skilling while leveraging AI's analytical strengths.

Closing reflection

Roundtable 1 highlighted that the central challenge of AI-supported negotiation is not whether artificial intelligence can be made more powerful, but whether it can be made more appropriate to the realities of negotiation practice. Current limitations were repeatedly framed not as transitional technical shortcomings, but as structural mismatches between how AI systems are designed and how negotiation actually functions, through ambiguity, relational judgment, institutional memory, and ethical responsibility.

The discussion underscored that progress will depend less on autonomous capabilities than on thoughtful system architecture: scaffolding reliability, preserving human agency, and embedding AI within institutional workflows rather than individual productivity tools. AI's analytical strengths are real and valuable, but they only translate into negotiation value when paired with human interpretation, accountability, and governance.

Ultimately, the roundtable reframed next-generation AI for negotiation as an infrastructural and institutional challenge rather than a purely technological one. Advancing the field will require designing systems that respect the asymmetry between human and machine capabilities, measure success beyond efficiency, and strengthen, rather than displace, the human judgment at the core of complex negotiation.

* * *

Report of Roundtable 2
Capacity Building and Empowering Negotiators
in the Age of AI

Building human and institutional capabilities for reliable AI-supported negotiation

Focus of the Roundtable

How can negotiators and organisations build the human, ethical, and institutional capacities required to integrate AI into negotiation practice without delegating judgment, responsibility, or accountability?

Moderator: **Francesco Marchi**, Director of Negotiation Expertise at ALTERNEGO, Lecturer, Sciences Po (Paris), College of Europe (Bruges) and University Paris 1 Sorbonne-Panthéon.

Contributors and themes

Leonardo Caporarello Professor of Practice of Negotiation and Leadership; Associate Dean for Online Learning, SDA Bocconi	*Negotiation in the Age of AI: The Human Leadership*
Vera Hampel Principal, Egger, Philips & Partner (Zurich)	*Human-Centered Capacity Building in AI-Enabled Negotiation Training*
Valon Murtezaj Professor of International Negotiation, IÉSEG School of Management	*The Use of AI in Diplomatic Negotiations*
Saloni Shukla United Nations Fellow and Emerging Leader 2025; Founder, Nova Ethos	*The Neoteric Diplomat in the Age of AI*
Jan Smolinski AI Researcher | Founder & CEO of Discurso.AI	*Who is the better negotiation judge, human or AI?*

Background

AI is already embedded in negotiation practice, largely through informal and individual experimentation rather than deliberate institutional strategy. Survey data indicate that nearly three quarters of practitioners now use AI daily or weekly, with usage intensifying as negotiators gain experience in prompting and interacting with models. At the same time, 97% of respondents explicitly called for more in-depth training, particularly on how to integrate AI into their everyday workflows and how to leverage its analytical and

sense-making capabilities to support strategic planning and scenario design. This growing reliance on AI is often occurring under pressure, including in live or high-stakes negotiations, while structured training pathways, organizational guidance, and governance frameworks remain limited or uneven.

Against this backdrop, the roundtable focused on capacity building and the design of training programs and courses capable of empowering both practitioners and their organisations to use AI in negotiation responsibly and effectively. Rather than concentrating on technical tool use alone, the discussion examined how training should strengthen human judgment, ethical responsibility, and institutional readiness, ensuring that AI enhances analytical depth and strategic foresight without displacing accountability or weakening organizational oversight.

Core themes emerging

Capacity building is primarily about human judgment in AI-human interaction, not technology.

A consistent through-line across the discussion was that **capacity building in the age of AI is not primarily a technical challenge**, but one of **human judgment, responsibility, and institutional readiness within human–AI interaction**. While AI is rapidly becoming embedded in professional practice, participants emphasized that decisions involving ethics, peace, protection, and accountability remain fundamentally human and cannot be delegated or automated.

AI reshapes negotiation into a socio-technical ecosystem.

Negotiation is no longer a purely human interaction but a hybrid environment where humans and AI systems operate in parallel. Contributors described a transition from human-led negotiation to AI-augmented practice, toward future hybrid socio-technical negotiation ecosystems, i.e. a negotiation environment in which human negotiators and AI systems operate together as an integrated system rather than as separate or sequential actors. In this projected future, AI will provide analytical capacity, memory, scenario generation, and feedback, while humans retain judgment, ethical responsibility, and decision authority.

The potential impact of AI on negotiation outcomes can be significant: such ecosystems can improve preparation quality, consistency, and institutional learning, while enabling negotiators to manage greater complexity and uncertainty in real time. At the same time, these augmented outcomes increasingly depend on how roles are distributed, the interaction governed, and accountability are designed, as poorly structured integration risks over-reliance on AI, erosion of judgment, and distorted incentives. These environments will require clear role boundaries, governance mechanisms, and sustained human oversight.

Relational and emotional dimensions remain structurally human.

While AI excels at speed, analytical structure, and pattern recognition, participants repeatedly emphasized that relational depth, contextual sense-making, legitimacy, and ethical interpretation remain human strengths. Research shared during the discussion suggested that AI can influence emotional regulation and engagement but not replace relational judgment. Participants agreed that while AI is widely trusted for its **analytical strengths**, it lags in terms of **relational understanding**, especially in negotiations characterized by power imbalance. Under such conditions, emotional awareness, interpretive skill, and relational intelligence become more important, not less.

Capacity Building Requires Institutional Systems, Not Only Skilled Individuals

A recurring concern was the widening mismatch between individual adoption and institutional preparedness. Many practitioners are experimenting with AI independently, while organisations lack capacity to review internal practices, shared standards, data governance, accountability frameworks, and institutional memory systems.

Approximately one-third of organisations represented were described as not yet meaningfully engaged with AI integration, while others are progressing unevenly. This imbalance shifts risk and responsibility onto individuals, weakens accountability, and reduces organizational visibility into how AI is actually being used. From a capacity-building perspective, participants emphasized that strengthening individual skills without parallel investment in institutional systems is insufficient.

Unequal access to AI risks amplifying power asymmetries.

Advanced AI tools, data access, and organizational capacity risk reinforcing inequalities not only between institutions and states, but also across regions and organisations, further complicating already asymmetric negotiation relationships. Actors with greater resources may be able to saturate negotiation spaces with AI-generated analysis, scenarios, and data, shaping agendas and framing discussions to their advantage. Without corresponding human and institutional capacity to assess, filter, and contextualize this information, other parties may face cognitive overload rather than empowerment. This asymmetry creates a risk of informational saturation that can slow, distort, or even paralyze multilateral negotiation processes, particularly where time, trust, and interpretive capacity are already under strain.

Capacity building therefore has a normative dimension: democratizing access and ensuring transparency is not only technical but ethical.

Key tensions

Efficiency versus agency.

Here again, participants expressed concerns about the loss of agency for the sake of productivity. AI demonstrably improves efficiency in preparation, analysis, and

feedback, including in training and assessment contexts where AI may outperform human evaluators. However, participants warned that efficiency gains can come at the cost of cognitive offloading and reduced human engagement, raising questions about when superior performance justifies such delegation.

Integration versus over-reliance.

Two models of human–AI interaction were identified:

- A "centaur" model, where tasks are clearly divided between humans and AI.
- A more fluid, real-time integration model, where boundaries blur.

While both models can increase efficiency and performance, several participants cautioned that deeper integration heightens the risk of over-reliance. As AI outputs become more sophisticated and persuasive, it becomes easier for humans to disengage critical judgment. While AI may optimize outcomes, it **cannot bear the consequences of its analysis and recommendations**. Judgment related to peace, ethics, protection, and responsibility cannot be outsourced. The future of diplomacy, several speakers argued, will be determined by **who remains in charge and accountable** as these systems evolve.

Operational and institutional implications

Human-led processes must be treated as a design principle, not a slogan. Capacity building should reinforce active human judgment rather than assume it will persist by default. This includes explicit checkpoints for decision-making, responsibility attribution, and ethical review.

Training should move beyond tool use to judgment formation.

Participants emphasized the need to train intuition, prioritization, ethical literacy, emotional awareness, and the ability to interpret AI outputs critically, including understanding how AI interacts with other technologies within broader ecosystems.

Institutions must develop ambidexterity.

Organisations need the capacity to experiment under uncertainty while also consolidating learning into shared standards, governance frameworks, and institutional memory. Individual skill development without parallel institutional investment was widely viewed as insufficient.

Governance frameworks must clarify responsibility.

Clear rules on data protection, confidentiality, explainability, and accountability are required to prevent responsibility from defaulting to individuals operating without institutional backing.

Open questions / areas for further inquiry

- How can institutions preserve and operationalize judgment in environments where AI consistently outperforms humans on specific evaluative tasks?

- What governance models are best suited for hybrid socio-technical negotiation systems operating across organizational and national boundaries?

- What forms of training prevent de-skilling while leveraging AI's strengths in learning, feedback, and preparation?

Closing reflection

Taken together, the roundtable suggested that empowering negotiators to use AI is **less about teaching tools** and more about reinforcing **human judgment, institutional responsibility, and ethical visibility**. The central question raised throughout was not whether AI can improve performance, but **who remains accountable as AI becomes embedded in negotiation practice**.

* * *

Report on Roundtable 3

Managing Cognitive, Ethical, and Institutional Risks

Safeguarding judgment, meaning, and accountability in AI-assisted negotiation

Focus of the Roundtable

How can negotiators and institutions identify, mitigate, and govern the cognitive, ethical, and institutional risks introduced by AI in negotiation and mediation, particularly in high-stakes, fragile, or politically sensitive contexts?

Moderator: **Annie Chaloux,** Associate professor, Université de Sherbrooke, holder of the Quebec Research Chair in Climate Scientific Diplomacy.

Contributors and themes

Ugne Fink-Jensen Managing Partner, Firm&Fair	*AI in Negotiations: Overreliance, De-Skilling, and the Semiotic Gap*
Emily Knowles Project Manager and Consultant UNOPS	*Applying Cultural Mapping to AI-Enabled Negotiation*
Abdel Rahman Alzorgan Public Information Assistant, Office of the UN Special Envoy for Yemen	*A Governance Framework for AI in Conflict Mediation: Where the Machine Stops*
Tea Mustać LL.M. Candidate, Harvard Law School; Head of Operations, Institute for Global Negotiation	*The Hidden Cost of Efficiency*
Dariha Erketaeva Negotiation Specialist FATC	*The Cognitive Costs and Benefits of AI for Frontline Negotiators*

Background

AI is now deeply embedded in negotiation and mediation practice, especially in preparation, analysis, drafting, and sense-making. Survey data and focus group discussions confirmed that this integration is accelerating, often driven by individual practitioners operating without clear organizational guidance. While AI-generated outputs frequently appear coherent, confident, and actionable, this surface quality masks significant risks related to bias, misinterpretation, and misplaced trust.

The roundtable examined how these risks materialize cognitively, ethically, and institutionally, and how they can undermine negotiation quality, legitimacy, and accountability if left unaddressed. Rather than focusing on technical safeguards alone, the discussion centred on how human judgment, cultural interpretation, and institutional governance must be deliberately designed into AI-assisted negotiation systems.

Core themes emerging

The semiotic gap is the most insidious risk.

A central and widely shared concern was the gap between AI linguistic fluency and actual understanding. AI systems can generate polished, persuasive text while failing to grasp intent, subtext, cultural nuance, or relational history. In negotiation, this gap between what is said, what is meant, and what is interpreted can lead to misreading trust signals, overlooking warnings, or mischaracterizing positions. Fluency increases the danger by concealing incomprehension, particularly in high-context or cross-cultural settings.

Cognitive offloading produces fragile negotiators.

Participants highlighted that delegating preparation, framing, and judgment to AI generates short-term efficiency but weakens long-term resilience. Evidence from neuroscience research indicates that reliance on generative AI for complex thinking reduces neural engagement, with cognitive adaptation persisting even after AI is removed. In negotiation terms, this manifests as better-looking preparation coupled with reduced adaptability, weaker situational awareness, and diminished judgment when negotiations move off script.

Over-reliance is reinforced by automation bias and organizational gaps.

AI outputs tend to be treated as authoritative, especially under time pressure or when baseline negotiation capacity is uneven. Automation bias, anchoring effects, and groupthink were repeatedly identified as mechanisms through which probabilistic AI models narrow rather than expands thinking. These risks are magnified when organisations provide tools without proper standards, training, or governance, effectively shifting responsibility onto individuals.

Cultural and contextual blindness remains a structural limitation.

LLMs are predominantly trained on dominant language patterns and struggle with dialects, indirect communication, pragmatic intent, and culturally embedded persuasion styles. The discussion emphasized that negotiation meaning is often carried by what is omitted, implied, or expressed non-verbally, dimensions that AI systems cannot reliably capture. This limitation is particularly consequential in multilateral, cross-cultural, and conflict-affected contexts.

Hybrid human–AI systems outperform AI-only approaches when designed properly.

Operational examples demonstrated that hybrid systems, where AI handles volume and pattern detection while humans retain responsibility for interpretation and judgment, can significantly outperform AI-only models. In mediation contexts, such systems compressed processing time while preserving accuracy through human verification layers. The key design principle repeatedly articulated was that AI should aggregate and surface signals, but humans must preserve meaning.

Key tensions

Speed versus safety.

AI compresses time and accelerates workflows, but also alters expectations around pace, potentially reducing patience and deliberation in negotiation processes. While speed can enhance inclusion and responsiveness, unchecked acceleration increases the risk of error, misinterpretation, and premature convergence.

Efficiency versus accountability.

Efficiency gains often obscure responsibility. When AI-generated outputs circulate without clear attribution, escalation protocols, or correction mechanisms, accountability becomes diffuse. High-profile failures across sectors illustrated how institutional responsibility persists even when errors originate from individual AI use.

Inclusion versus distortion.

AI can expand participation in negotiation and mediation processes by aggregating and processing large volumes of input from diverse actors who would otherwise remain unheard. This capacity creates new opportunities for inclusion, particularly in multilateral, geographically dispersed, or conflict-affected contexts. However, inclusion at scale introduces a parallel risk of distortion. Without robust governance and human verification, AI-mediated aggregation may flatten minority perspectives, misinterpret culturally coded language, or amplify dominant narratives embedded in training data. Translation errors, loss of contextual nuance, and automated sentiment analysis can further transform meaning, creating a false sense of consensus or misrepresenting dissent.

As a result, increased participation does not automatically translate into more legitimate or accurate representation; inclusion only strengthens negotiation outcomes when institutions retain the capacity to interpret, validate, and contextualize AI-processed inputs rather than treating them as neutral or self-evident.

Operational and institutional implications

Context firewalls are essential.

Participants emphasized the importance of establishing explicit boundaries, "context firewalls", that define where AI assistance must stop, and human judgment must intervene. In high-stakes negotiations, certain domains are particularly sensitive to misinterpretation, including political sentiment analysis, cultural and linguistic nuance, assessments of trust and credibility, and the interpretation of implicit signals. In these areas, AI-generated outputs may appear coherent while masking contextual errors or false inferences, creating disproportionate risk. Mandatory human verification was therefore identified as a non-negotiable safeguard, ensuring that meaning is interpreted by actors with situational, cultural, and relational awareness.

Context firewalls do not aim to limit the usefulness of AI, but to channel it appropriately: AI can aggregate information, surface patterns, and flag anomalies, while humans retain exclusive authority over interpretation, judgment, and decision-making. By clearly demarcating these boundaries, institutions can benefit from AI's speed and scale without allowing automated systems to autonomously define meaning or steer negotiation outcomes in domains where errors are difficult to detect and costly to correct.

Governance must precede scale.

Institutions should not deploy AI at scale before establishing clear governance frameworks that define who can use AI systems, for what purposes, and under which conditions. This includes explicit rules on data access and boundaries, standards for verification and human review, escalation procedures for errors or high-risk outputs, and requirements for transparency toward internal and external stakeholders. Scaling AI without these safeguards risks normalizing unverified outputs, diffusing responsibility, and embedding errors into routine practice.

Participants emphasized that trust in AI-assisted negotiation depends not on the absence of mistakes, but on visible mechanisms for detecting, documenting, and correcting them. Logging errors, acknowledging failures, correcting outputs when they circulate publicly, and disclosing when and how AI is used were identified as essential practices to preserve institutional credibility, accountability, and confidence as AI becomes more deeply integrated into negotiation processes.

Training must address cognitive risk, not only tool use.

Capacity building efforts should move beyond familiarizing practitioners with AI functionalities and instead explicitly address the cognitive risks associated with sustained AI use in negotiation. These include cognitive offloading, automation bias, and over-reliance, as well as the gradual erosion of executive functions that are critical under pressure, such as attention control, emotional regulation, situational awareness, adaptability, and real-time judgment. Evidence discussed during the roundtable

underscored that while AI can enhance short-term efficiency and confidence, it may also weaken neural engagement and reduce cognitive resilience when practitioners routinely delegate core thinking tasks to automated systems.

Participants emphasized that without deliberate training design, AI can quietly displace the mental reflexes negotiators rely on when negotiations move off script, introducing fragility precisely at moments that demand human judgment most. Effective training should therefore cultivate reflective and sequenced use of AI, encouraging practitioners to engage in initial framing, interpretation, and decision-making before using AI to test, enrich, or challenge their analysis. Training should also address how AI-generated fluency can mask errors or contextual gaps, reinforcing the need for critical interrogation rather than passive consumption of outputs. The objective is not to limit AI use, but to regulate what is cognitively offloaded and what must remain human, ensuring that AI integration strengthens, rather than weakens, judgment, resilience, and accountability in negotiation practice.

Institutional responsibility cannot be outsourced.

While individual practitioners are the primary users of AI systems, institutions remain ultimately accountable for the outcomes those systems influence. Participants repeatedly identified the practice of providing AI tools without accompanying guidance, standards, or oversight as a systemic failure that amplifies risk rather than mitigating it.

In the absence of institutional frameworks, responsibility is implicitly shifted onto individuals operating under pressure, often without clarity on acceptable use, escalation pathways, or liability. This not only weakens accountability but also reduces organizational visibility into how AI is actually shaping decisions and behaviours. Effective governance therefore requires institutions to assume responsibility for defining boundaries, setting expectations, monitoring use, and responding to failures, recognizing that accountability for AI-assisted negotiation cannot be delegated to individual discretion alone.

Open questions / areas for further inquiry

- How can institutions design training that preserves human judgment while benefiting from AI-driven efficiency?

- What governance models are most effective for managing AI use in multilingual, high-context negotiation environments?

- How should organisations define and enforce "stop rules" for AI in live or high-stakes negotiations?

- How can transparency about AI use be operationalized without undermining trust or legitimacy?

Closing reflection

The roundtable converged on a clear conclusion: the most serious risks of AI in negotiation are not technical failures, but cognitive, interpretive, and institutional ones. AI changes how negotiators think, how fast they move, and how meaning is constructed. Managing these risks requires deliberate system design that keeps humans at the centre of judgment, preserves accountability, and treats governance as a prerequisite rather than an afterthought.

* * *

How artificial intelligence reshapes narrative power, reasoning, and world-making

Focus of the Roundtable

How does the integration of AI into negotiation, diplomacy, and peacebuilding reshape meaning-making, narrative authority, and legitimacy, and what risks does this pose for high-stakes, culturally embedded, and politically sensitive negotiations?

Moderator: **Delphine Alles**, Professor of Political Sciences, Vice-President INALCO, Scientific Director DÉCRIPT

Contributors and themes

Charles Tenenbaum Associate Professor of Political Science, Sciences Po Lille **Iris Poiré Hernandez** Dual Master's Student, Sciences Po & Centrale Lille	*Algorithms and the Shaping of Complex Negotiation Processes: The Case of Digital Peacebuilding*
Bahar Sercin Halefoglu CEO & Founder, ALSE Data; AI Policy & Diplomacy Practitioner	*Reframing AI's Role in Institutional Negotiation*
Louise Beaumais Postdoctoral Researcher, INALCO / CESSMA	*Beyond Technological Optimism: Indigenous' Peoples Contestations of AI*
Martin Wählisch Associate Professor, University of Birmingham	*Reasoning with Machines? The Power and Pitfalls of AI in High-Stakes Negotiations*

Background

Negotiation and diplomacy are not only processes of decision-making, but practices of meaning-making through which realities, identities, and futures are constructed. As AI systems become increasingly embedded in context analysis, briefing, sense-making, and scenario exploration, they do more than accelerate analysis: they shape which narratives are visible, which histories are foregrounded, and which interpretations appear reasonable or authoritative.

Survey results and practitioner discussions confirmed widespread reliance on AI for contextualization, preparation of briefs, and exploration of negotiation pathways. While most practitioners view AI as a decision-support tool rather than a decision-maker, the roundtable addressed a deeper concern: that algorithmic systems increasingly participate in defining problems, framing options, and justifying outcomes, thereby reshaping the conditions under which negotiations unfold.

Core themes emerging

AI reshapes narrative authority and problem definition.

Algorithms do not merely process information; they structure visibility and relevance by filtering, ranking, and reformulating inputs. In doing so, they pre-shape how negotiation problems are defined, and which solutions appear plausible. This raise concerns that AI may silently displace narrative authority, especially in contexts where historical memory, cultural interpretation, and symbolic meaning are central.

Rationalization risks oversimplifying conflict dynamics.

AI systems tend to approach conflict as a problem to be modelled, optimized, and managed through identifiable variables and predicted outcomes. AI can provide valuable analytical clarity by organizing complex information, highlighting patterns, and surfacing potential leverage points. However, participants stressed that peace processes cannot be reduced to rational optimization without losing essential dimensions of how conflicts actually evolve and are resolved. Emotions, trauma, symbolic acts, misperception, and shifting identities often shape decisions more powerfully than calculated interests, particularly in protracted or violent conflicts.

Negotiation breakthroughs frequently emerge from ambiguity, informal gestures, or moments of apparent irrationality that escape formal modelling. Excessive reliance on AI-driven rationalization therefore risks flattening the lived realities of conflict, privileging what is measurable over what is meaningful, and narrowing the space for creative, adaptive, or politically sensitive solutions that depend on human judgment rather than algorithmic coherence.

Meaning-making and world-making are inseparable.

Negotiations do not merely resolve disputes or allocate resources; they actively construct shared understandings of reality by defining problems, identities, responsibilities, and futures. Through this process, language, framing, and interpretation shape not only what agreements contain, but how they are perceived, accepted, and remembered. By influencing how information is selected, prioritized, summarized, and justified, AI systems increasingly participate in this meaning-making process.

This influence extends beyond technical support to the production of narratives that render certain outcomes reasonable or legitimate while marginalizing others. In

peacebuilding contexts, where agreements must be socially appropriated by affected communities to endure, shifts in narrative authority carry profound implications. If AI-mediated framing subtly reshapes how conflicts and solutions are understood, it can affect ownership, trust, and long-term legitimacy, underscoring the need for careful governance of AI's role in shaping not only decisions, but the worlds those decisions help to create.

Opacity undermines legitimacy.

Most large language models operate as "black boxes," offering outputs whose internal logic, data provenance, and weighting of sources remain largely inaccessible to users. This limits negotiators' ability to explain how conclusions were reached, why certain interpretations prevailed, or which assumptions shaped the analysis.

Such opacity can conceal dominant worldviews, embedded normative choices, and structural power asymmetries that quietly influence outcomes. Participants warned that reliance on opaque systems risks weakening the legitimacy of negotiation processes and agreements, particularly in contexts where accountability, transparency, and trust are central to acceptance.

When negotiators cannot meaningfully justify how AI-informed positions were developed, they may struggle to defend decisions to stakeholders, affected communities, or oversight bodies. In democratic or peacebuilding settings, where legitimacy depends on the ability to contest, revise, and own outcomes, opacity transforms AI from a support tool into a source of fragility, underscoring the need for explainability, documentation, and human interpretive authority.

Knowledge asymmetries and representation

AI amplifies existing power asymmetries.

AI systems privilege knowledge that is already digitized, standardized, and available in dominant languages. This systematically advantages well-resourced actors and marginalizes perspectives that are oral, contextual, or culturally embedded. The discussion highlighted how unequal access to data and computational capacity can intensify asymmetries between parties.

Indigenous critiques expose structural limits of AI.

Indigenous peoples' contestations of AI illustrate broader risks of exclusion, misrepresentation, and epistemic erasure. Indigenous knowledge often resists extraction into machine-readable corpora and loses meaning when decontextualized or translated into dominant categories. Participants emphasized that AI can inadvertently perpetuate colonial logics by fixing knowledge, reinforcing extractive practices, and privileging certain worldviews over others.

Inclusion without interpretation risks distortion.

While AI can scale participation and aggregate large volumes of input, this does not guarantee accurate representation. Without human interpretation, AI-mediated summaries may flatten minority views, misread culturally coded language, or falsely signal consensus, undermining rather than enhancing legitimacy.

Reasoning, authority, and responsibility

From tools to authorities.

A central concern raised during the roundtable was the gradual shift of AI from analytical infrastructure to systems that actively recommend, rank, and justify courses of action. As AI outputs move upstream, from organizing information to proposing options and providing rationales, they risk being treated not merely as inputs to human deliberation, but as authoritative reasons for decision-making. This subtly reconfigures who or what defines what is considered "reasonable," acceptable, or inevitable in a negotiation.

AI systems operate under uncertainty but speak with confidence, producing artificial clarity that can crowd out alternative interpretations. This can reduce contestability, discourage dissent, and make non-AI-endorsed options appear reckless or illegitimate. Participants stressed that the risk is not primarily error, but premature closure. In high-stakes contexts, qualities such as confidence, coherence, and speed are often interpreted as markers of rationality and competence, qualities that AI systems consistently project. As a result, AI-generated recommendations may acquire disproportionate persuasive power, narrowing deliberation and shifting authority away from human judgment toward algorithmic reasoning, even when responsibility for outcomes formally remains with negotiators.

Governance and design implications

Clear boundaries for AI reasoning.

Participants emphasized the need to define where AI reasoning must stop. While AI can explore options, stress-test assumptions, and surface blind spots, it should not autonomously frame negotiation parameters, define acceptable trade-offs, or justify final outcomes.

Hard stops and procedural safeguards.

Participants emphasized the need to translate principles of human oversight into concrete procedural safeguards that actively structure how AI is used in negotiation processes. Proposed measures included mandatory human decision pauses that interrupt automated workflows and require deliberate reflection before action is taken,

particularly in high-stakes moments. Several speakers stressed the importance of requiring AI-generated analyses or justifications to be rewritten by humans in their own words before being relied upon or entered into the official record, reinforcing ownership and accountability. Together, these measures are designed to prevent premature closure, preserve deliberation, and ensure that AI remains a support for human reasoning rather than a substitute for it.

Open questions

- How can negotiators preserve interpretive pluralism and ambiguity while benefiting from AI-enabled efficiency?

- Who controls the narratives embedded in training data, and how can marginalized worldviews meaningfully shape AI-assisted negotiation processes?

- How can legitimacy be sustained when AI systems increasingly participate in justification and sense-making?

- What institutional arrangements are required to prevent authority leakage while enabling analytical support?

Closing reflection

Round Table 4 underscored that AI's most profound impact on negotiation lies not in automation, but in its influence over meaning, authority, and legitimacy. Negotiation is not merely about choosing among options, but about constructing shared realities under conditions of uncertainty, power asymmetry, and historical depth. If AI is to support negotiation rather than simplify it away, it must be governed not only as a technical tool, but as a participant in meaning-making, one whose role, limits, and assumptions are made explicit, contestable, and subordinate to human judgment and responsibility.

* * *

Section IV: Overarching Observations from the Forum Process

Across the survey, focus groups, and roundtable deliberations, a coherent and consistent picture emerges of how artificial intelligence is currently reshaping negotiation, mediation, and diplomacy, and of the gaps that must be addressed before its use can be considered safe, effective, and legitimate.

1. **AI is already deeply embedded in negotiation practice, but in an informal, uneven, and weakly governed manner.**

Survey data show widespread and regular use of AI by mid- to senior-level negotiators, primarily for preparation, analysis, drafting, and sense-making. Focus group discussions confirm that this use is largely self-directed, driven by operational necessity rather than institutional strategy. Roundtables reinforced that the question is no longer whether AI will be used, but how its use can be made visible, structured, and accountable. The pace of individual adoption has clearly outstripped institutional readiness.

2. **Practitioners experience real and significant benefits but remain acutely aware of AI's limitations.**

Across all stages of the Forum, participants consistently emphasized gains in efficiency, analytical breadth, and cognitive support. At the same time, there is a strong and shared recognition that AI performs poorly on the relational, ethical, cultural, and political dimensions that are central to negotiation. This duality, high utility combined with structural inadequacy, explains why practitioners use AI intensively while simultaneously resisting its elevation to an autonomous or authoritative role.

3. **Human judgment is not being displaced, but it is under pressure.**

Survey respondents report strong verification behaviours and a clear sense of personal responsibility for AI-assisted outcomes. However, focus group and roundtable discussions highlight growing risks of cognitive offloading, automation bias, and false confidence, particularly under time pressure and in high-stakes contexts. The risk is not immediate replacement of human judgment, but its gradual erosion if AI integration is not deliberately designed to preserve reflection, contestation, and accountability.

4. **The main risks of using AI in negotiation are cognitive, ethical, and institutional rather than purely technical.**

While technical issues such as hallucinations and data gaps remain important, the Forum consistently surfaced deeper vulnerabilities: misplaced trust in opaque systems, bias embedded in training data, loss of contextual sensitivity, and diffusion of accountability within organisations. Institutional gaps, lack of training, governance, and shared standards, amplify these risks by shifting responsibility onto individual users operating without adequate support.

5. **Current AI systems are misaligned with how negotiation works.**

Roundtable discussions underscored a structural mismatch between AI architectures optimized for coherence, speed, and plausibility, and negotiation processes that depend on ambiguity, informal signalling, relational judgment, and ethical restraint. This mismatch limits trust in AI outputs and constrains their appropriate use. The challenge ahead is therefore less about making AI more powerful, and more about making it more appropriate to negotiation as a human, social, and political practice.

6. **Capacity building must focus on human judgment and institutional integration, not just tools.**

A strong convergence emerged around the view that training negotiators to "use AI" is insufficient. What is needed are capabilities to decide when to use AI, when not to, how to interpret and challenge outputs, and how to manage confidentiality, bias, and accountability. Equally important is institutional capacity: governance frameworks, shared standards, and organizational memory systems that prevent AI-related risk from being individualized.

Discussions on impact revealed that AI increasingly participates in framing problems, structuring narratives, and defining what appears reasonable or authoritative. While AI can democratize access to information, unequal access to models, data, and infrastructure risks reinforcing existing power asymmetries. Without transparency and negotiated safeguards, AI use risks undermining trust, legitimacy, and ownership of negotiation outcomes.

Taken together, these observations point to a central conclusion: **the integration of AI into negotiation is not primarily a technological transition, but a socio-technical and institutional one.** The distinguishing contribution of the AI Negotiation Forum lies in making this reality visible, by grounding analysis in practitioner experience and linking observed practice to research, training, and governance debates. These observations provide the foundation for identifying priorities and next steps in shaping responsible, human-centred AI-supported negotiation.

Conclusion and next steps

The first edition of the AI Negotiation Forum confirms both the depth of transformation already underway and the urgency of shaping it deliberately. Across the survey, focus groups, and roundtables, practitioners, experts, and researchers converged on a clear message: AI is becoming an integral part of negotiation practice, but its value depends on how it is governed, learned, and embedded within human judgment and institutional responsibility. The Forum's practice-led approach, grounding reflection and analysis in the lived experience of negotiators, constitutes its central contribution and provides a solid foundation for future work.

This process would not have been possible without the collective engagement of practitioners who shared their experience openly, the experts and researchers who brought critical perspective to the discussions, and the facilitators who ensured that dialogue remained anchored in practice. Special acknowledgment is due to the mid-career students at the Harvard Kennedy School, whose role in facilitating the focus groups and supporting the Forum's analytical process was essential. Their contribution exemplifies the Forum's broader ambition: to connect generations of practitioners, researchers, and educators in shaping responsible, human-centred uses of AI in negotiation, mediation, and diplomacy.

Next Steps and Action Plan

The findings of the AI Negotiation Forum point to an urgent need for coordinated action across multiple communities. AI is already shaping negotiation practice, but its integration remains uneven, weakly governed, and largely driven by individual initiative rather than collective strategy. The following action plan outlines priority next steps for four key audiences, reflecting their distinct responsibilities and leverage points.

1. **Practitioners: Strengthening Individual Agency and Collective Practice**

Practitioners are at the core of the Forum and the primary drivers of AI use in negotiation. The first set of actions concerns how negotiators, mediators, and diplomats can strengthen their own agency while reducing personal and institutional risk.

- A priority is the development of AI literacy beyond basic tool use. Practitioners are encouraged to move from tactical and reactive use of AI toward a more strategic and reflective integration. This includes developing the ability to decide when AI adds value, when it introduces risk, and when it should not be used at all, particularly in trust-sensitive, relational, or ethically charged situations.

- Practitioners should also actively exchange practices with peers, using professional networks, communities of practice, and structured forums to compare use cases, failure modes, safeguards, and emerging norms. Peer

learning is essential to counter isolation, reduce uneven practices, and surface collective standards before they are imposed externally.

- At the same time, practitioners are encouraged to engage their organisations proactively, advocating for clearer frameworks, guidance, and governance for AI use. This includes pushing for explicit rules on confidentiality, accountability, and disclosure, both internally and in interactions with other negotiation parties. Individual competence alone cannot compensate for institutional silence or ambiguity.

2. Institutions and Organisations: Closing the Capability and Governance Gap

For institutions, the Forum highlights a critical and widening gap between staff capabilities and organizational readiness. Addressing this gap is no longer optional.

Institutions should recognize the strategic cost of inaction. Organisations that fail to develop AI capabilities risk losing comparative advantage in analysis, preparation, and strategic foresight relative to peers and counterparts. At the same time, they expose themselves to significant legal, ethical, and reputational risks, as staff increasingly use AI tools in unregulated and opaque ways.

- A first priority is improving AI literacy at senior management level. Leaders must understand not only what AI can do, but how it reshapes judgment, accountability, risk distribution, and power dynamics. Without informed leadership, governance efforts will remain superficial or purely defensive.

- Institutions should also build internal capacity rather than relying solely on external tools. This includes developing internal training programs, monitoring and documenting AI use, gathering best practices across teams, and creating shared standards and review mechanisms. Governance frameworks should clarify acceptable use, escalation procedures, data protection requirements, and responsibility for AI-assisted decisions.

- Last but not least, institutions must ensure that AI augments rather than erodes professional capability. Training and oversight should be designed to preserve human judgment, creativity, and relational skills, preventing long-term de-skilling and over-reliance. The objective is institutional learning and resilience; not short-term efficiency gains alone.

3. Researchers and Experts: Advancing Tools, Methods, and Critical Insight

Researchers and experts play a central role in accompanying this transition and translating practitioner needs into innovation.

- A key next step is the development of tools, methods, and system architectures that are appropriate to negotiation as a human, social, and political practice.

This includes AI systems designed for human–AI collaboration, transparency, and bounded use, rather than autonomous optimization.

- Researchers are also encouraged to develop training models and evaluative frameworks that help practitioners challenge AI outputs, identify failure modes, and understand cognitive and ethical risks. Much remains to be learned about how AI affects judgment, trust, learning, and decision-making over time, particularly in high-stakes and cross-cultural contexts.

- Finally, the Forum points to a significant opportunity: when properly designed and governed, AI could meaningfully enhance negotiation, mediation, and diplomacy in addressing today's major challenges. Understanding how AI can contribute, without distorting meaning, legitimacy, or accountability, remains an open and essential research frontier.

4. Academic Institutions and Training Centres: Shaping the Next Generation of Practice

Academic institutions and training centres have a distinctive responsibility to anchor these developments in education and professional formation.

- A priority is to integrate AI into negotiation and diplomacy curricula in ways that go beyond technical familiarity. Training should focus on judgment, ethics, governance, and human–AI interaction, preparing practitioners to operate in hybrid socio-technical environments.

- Universities and training centres are also well placed to serve as neutral spaces for experimentation, reflection, and standard setting, bringing together practitioners, researchers, and technologists. Through simulations, labs, and applied research, they can test assumptions, surface risks, and contribute to shared norms before practices harden in the field.

- Finally, academic institutions can help ensure that AI integration strengthens, rather than undermines, the human foundations of negotiation, by reinforcing critical thinking, ethical reasoning, and reflective practice as core professional competencies in the age of AI.

* * *

Annex: Analysis of the AI Negotiation Forum Survey

The AI Negotiation Forum Survey examines how negotiators, mediators, and diplomats are using and experiencing artificial intelligence (AI) in their professional practice. Conducted in preparation for the AI Negotiation Forum, the survey aims to:

- Map the current use of AI across complex negotiation and mediation environments.
- Assess perceptions, confidence, and trust in AI among practitioners.
- Identify training needs for the safe and responsible integration of AI in negotiation and diplomacy.

The findings will directly inform the Focus Group Discussion on 10 January 2026 and will support the work of the Thematic Roundtables on 11 January, which will address AI research, capacity building for negotiation and mediation, and approaches to responsible innovation.

The Survey, composed of 28 questions, was launched on Typeform on 19 November and remained open until 20 December 2025. This analysis reflects the 483 responses received.

This preliminary analysis reviews:

1. The profiles of the respondents
2. The integration of AI in their work
3. Their perception of the impact of AI in their work
4. Concerns and expectations of participants toward the role of AI in negotiation and diplomacy
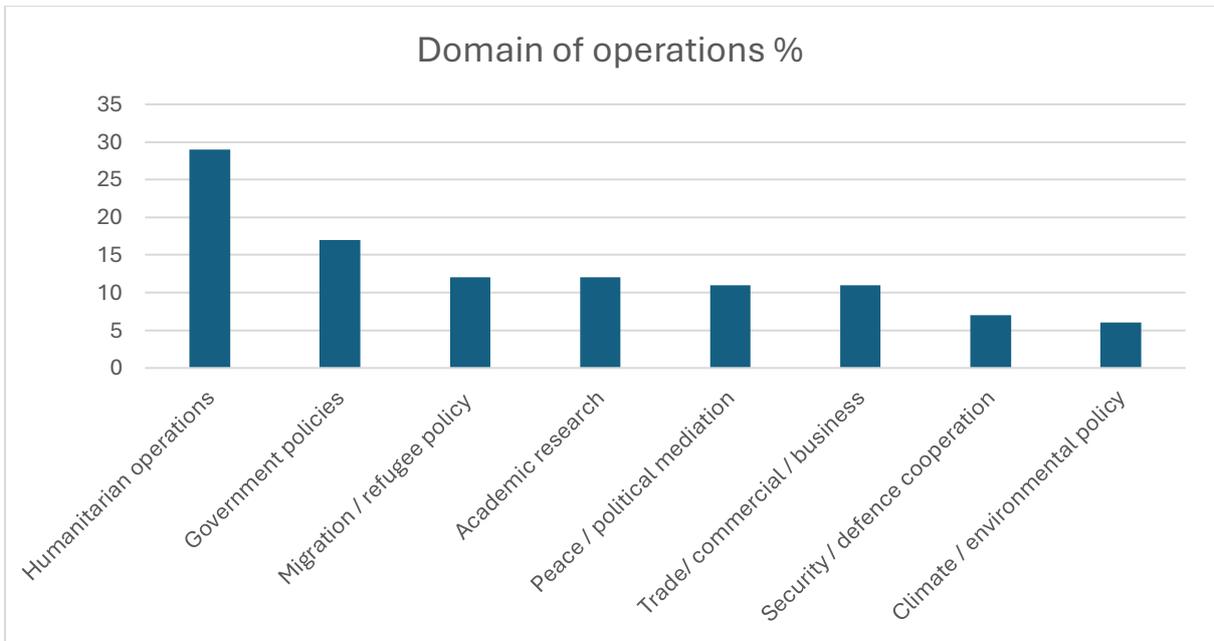
1. Profiles of the respondents

The Survey was distributed to approximately 9,000 potential respondents. A completion rate of around 5% aligns with expectations for a substantive cold survey of this kind.

Global and multi-regional practitioners form the largest cluster of respondents (~40–45%). This shows the survey attracted professionals engaged in cross-border negotiation environments typical of large multilateral, humanitarian, and diplomatic organisations.

Professional Roles

| Role	Value
Senior Executive/ Manager / Team Leader	~118
Field / Programme Officer	~78
Academic/ Researcher / Graduate Student	~66
Consultant / Independent expert	~52
Professional Negotiator / Mediator	~50
Government Representative/ Diplomat	~34
Policy Advisor/ Analyist	~28
Lawyer / Legal Advisor	~16

Respondents occupy a wide range of negotiation-relevant roles, with a strong concentration of senior managers, field and programme officers, professional mediators, consultants, policy advisors, and analysts.

Domain of operations %

Domain	Value
Humanitarian operations	29
Government policies	17
Migration / refugee policy	12
Academic research	12
Peace / political mediation	11
Trade/ commercial / business	11
Security / defence cooperation	7
Climate / environmental policy	6

This mix of domains reflects a practitioner community that is both operational and strategic: individuals who negotiate regularly in humanitarian access, public policy, migration, peace processes, organisational disputes, and multistakeholder coordination.

Their seniority and experience, most with more than eight years in negotiation, give the survey a high level of professional depth.
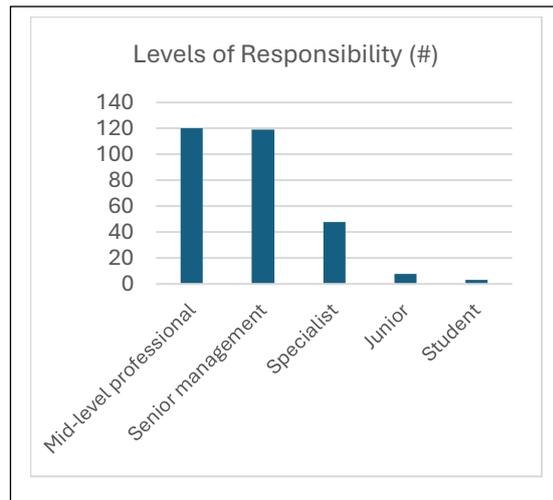
Respondents come primarily from international and regional organisations and international NGOs, with additional representation from academic institutions, civil society groups, private companies, and government agencies. This organisational landscape mirrors the ecosystem where complex negotiations typically take place and indicates that respondents bring perspectives from both global headquarters and field environments.
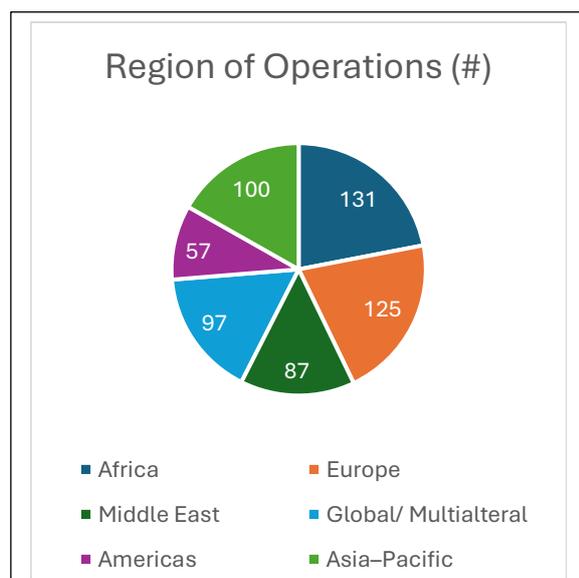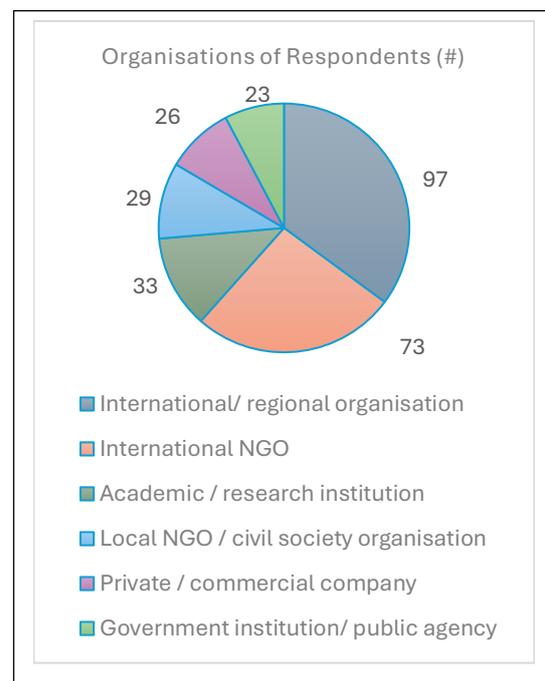

Figure 1Distribution per region of operations (some respondents mentionned more than one region)

Africa and the Middle East are the dominant regions of negotiation-intensive humanitarian and political contexts, together appearing in over half of responses. Europe is also strongly represented, primarily through institutional, headquarters, diplomatic, and research roles. Asia and the Americas appear as secondary clusters, each contributing important but smaller shares.

2. The integration of AI in their work

➡ Respondents are becoming increasingly familiar with AI tools

Levels of familiarity seems to increase steadily. Approximately one-third of respondents use AI daily or several times a week, and the majority report regular engagement for multiple tasks.

Practitioners rely on AI because it is useful, even as they remain cautious and aware of limitations. This combination of regular use, moderate confidence, and strong verification behaviours characterises a field that is moving quickly into applied AI practice while still building foundational literacy and institutional support.
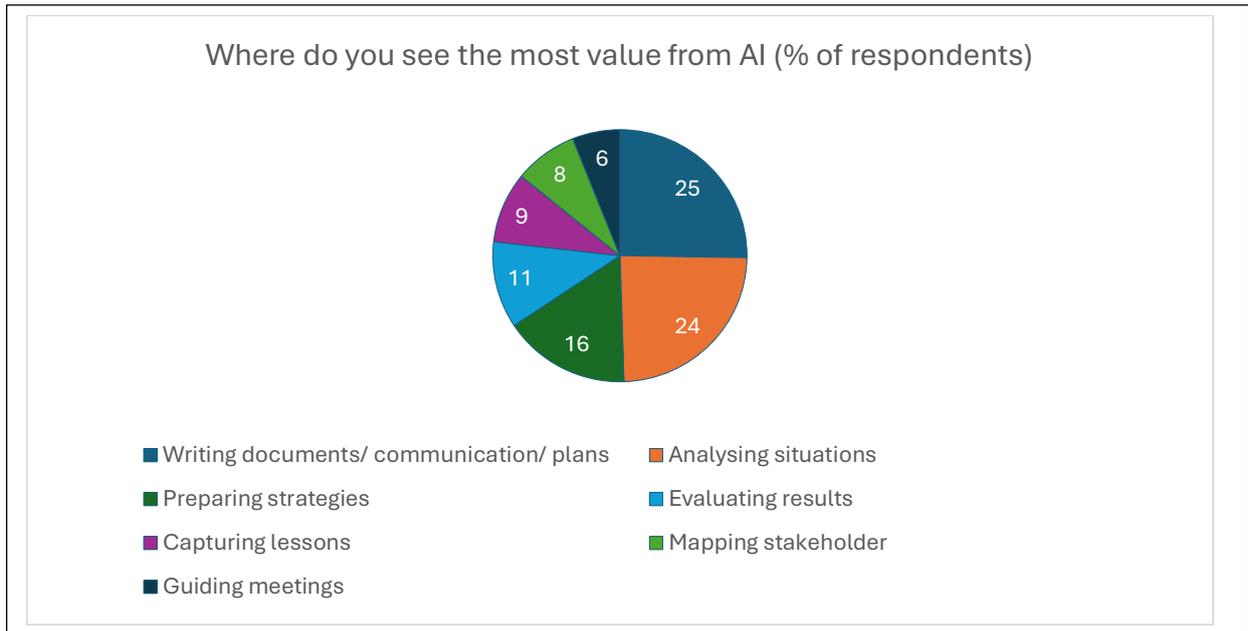
Frequency of AI use (# of respondents)
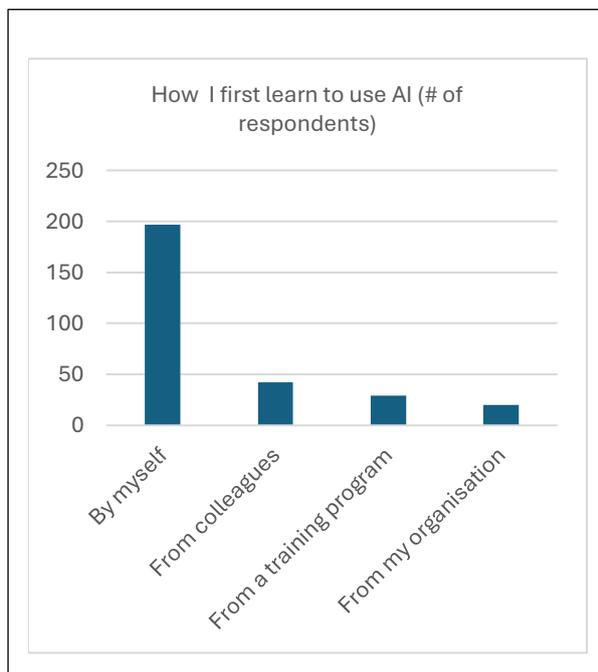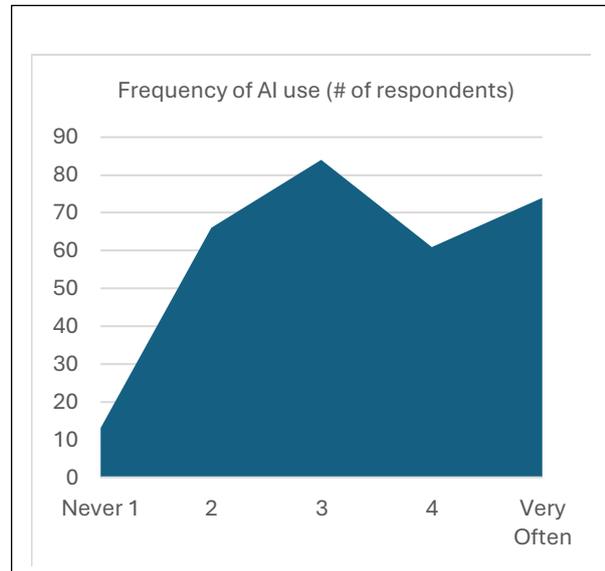
➡ People learn how to use AI mostly by themselves

Learning remains largely self-directed. Most respondents describe developing skills through experimentation, online resources, peer exchange, or task-specific problem-solving rather than through structured institutional training. This is reinforced by the strong demand for capacity-building and technical support: training and hands-on guidance are the most requested forms of institutional assistance.

How I first learn to use AI (# of respondents)

In terms of tools, OpenAI's ChatGPT is by far the most widely used (preferred by 43% of respondents), followed by Microsoft Copilot (21%), Google Gemini (16%), and smaller clusters using Claude (7%) and Perplexity (7%). All other tools (Mistral, DeepSeek etc.) amounts to 6% only.

AI is already reshaping core elements of negotiation work

Where do you see the most value from AI (% of respondents)

- Writing documents/ communication/ plans: 25
- Analysing situations: 24
- Preparing strategies: 16
- Evaluating results: 11
- Capturing lessons: 9
- Mapping stakeholder: 8
- Guiding meetings: 6

Respondents most frequently use AI for:

- Summarising information and organising discussions
- Drafting documents, preparing briefs, and generating talking points
- Analysing contexts, patterns, trends, and risks
- Exploring strategic options and scenario pathways

This indicates broad reliance on AI for sense-making, preparation, and communication.

3. *Their perception of the impact of AI in their work*

Perceived benefits are strong, especially regarding efficiency and cognitive support

Respondents offered rich and detailed accounts of how AI has reshaped their approach to work, with nearly 80 percent sharing a concrete moment when its use led to a noticeable shift in their professional practice.

The strongest pattern relates to gains in:

- **Efficiency and speed**, where AI enabled users to meet tight deadlines, automate labour-intensive tasks, and redesign their workflows around faster drafting and information processing.

- **Analytical depth and decision support**: many described instances in which AI revealed trends they had not seen, helped structure arguments, or provided alternative perspectives that broadened their thinking.

- **Writing, drafting, and communication support**, with respondents noting that AI now often serves as the starting point for reports, briefs, talking points, and presentations.

- **Learning and upskilling**, using AI as a tutor or explainer to acquire new competencies more quickly.

Others described how AI sharpened their **risk awareness**, exposing hallucinations and biases that reinforced the need for verification and critical judgment. A smaller but meaningful segment referenced **strategic and negotiation preparation**, where AI helped simulate roles, explore scenarios, or map stakeholders.

➡️ **Verification and human judgment remain central**

Respondents show very strong agreement on (average of 1 to 5):

- Double-checking outputs (4.2)
- Using human intuition alongside AI (4)
- Feeling personally responsible for AI-assisted outcomes (3.9)

This demonstrates a high verification culture and indicates that AI is used as a support tool, not a decision-maker.

4. *Concerns and expectations of participants toward the role of AI in negotiation and diplomacy*

➡️ **Risks are widely recognised and multifaceted**

Strongest concerns include (average of 1 to 5):

- Confidentiality and data sensitivity risks (3.8)
- AI bias and misinterpretation (3.8)
- Loss of creativity or reflection (3.8)

Additional risks (135 responses) reinforce four areas:

- Cognitive risks – over-reliance, loss of critical thinking
- Ethical risks – bias, misinformation, hallucinations
- Systemic risks – power imbalances, lack of transparency, dependency
- Operational risks – misinterpretation, missing context, poor judgement

Most respondents have experienced inaccuracies (combined yes + narratives = 71%).

➡ Trust in AI is growing but remains limited

- Respondents are ready to adjust their approach when AI offers new perspectives (3.5)
- However, this trust grows slowly over time (2.9)
- Peer influence is low (2.7)

Trust in AI is pragmatic, not relational: grounded in perceived utility and the ability to override AI, not social norms.
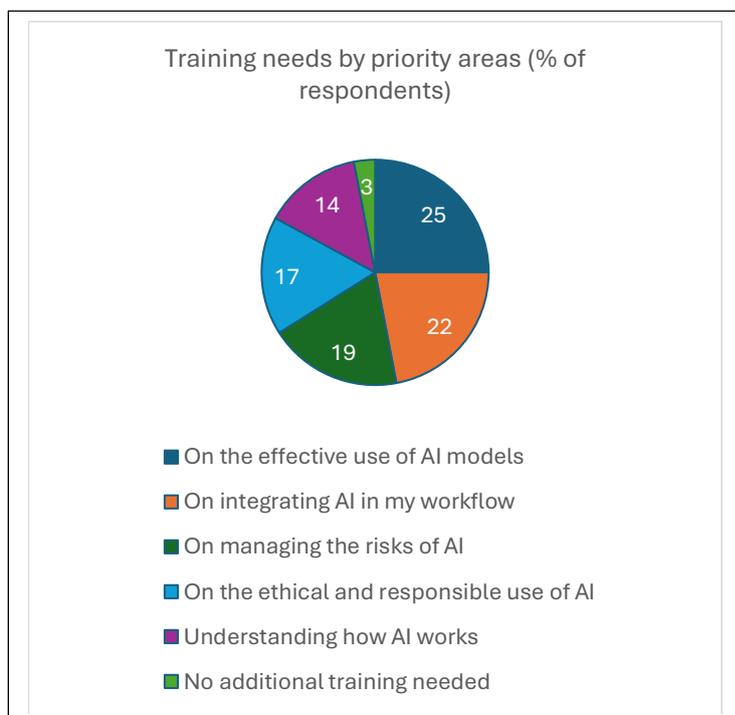
➡ Institutions are not keeping pace with individual use of AI by their staff

Institutional readiness scores are significantly lower:
- Leadership encouragement (2.8)
- Access to internal training (2.7)
- Peer learning (2.5)
- Workflow integration (2.4)

Organisations seems to lag behind staff who are already using AI intensively. This widens the capability and governance gap.

➡ Respondents overwhelmingly request training and technical support

Training needs by priority areas (% of respondents)

- On the effective use of AI models
- On integrating AI in my workflow
- On managing the risks of AI
- On the ethical and responsible use of AI
- Understanding how AI works
- No additional training needed

There is strong demand for practical skills. Effective use of AI models and integrating AI in workflow account for 47% of all training needs, signalling practitioners want applied, hands-on instruction.

Risk management is a major concern (19% of selections) emphasize managing AI risks, aligned with concerns over confidentiality, bias, and misinformation. Ethics remains a core expectation (16.5%), reflecting institutional and political sensitivities around AI. Foundational literacy is needed but secondary. Only 13.8% of respondents selected understanding how AI works, showing a priority on what to do with AI over how AI functions technically. Minimal number report "no training needed" (3%) indicating broad recognition of training needs across the field.

➡ Widespread ambivalence about AI in multilateral diplomacy

When asked about AI as a formal participant in multilateral processes:

- 49% offer mixed/nuanced reflections on such development
- 47% highlight risks
- 3% only highlight opportunities

Concerns focus on:

- Undermining legitimacy of the actors by inserting AI automation
- Transparency issues in using AI models in multilateral exchanges
- Potential political bias
- Job displacement
- Over-reliance and loss of human judgment
- Equity of access between states

This ambivalence underscores the need for prudence in the deployment of AI in multilateral circles, proper governance, and safeguards.

➡ Views on the future of AI in negotiation are balanced but cautious

Across all responses, a consistent picture emerges:

- Practitioners are using AI extensively to enhance negotiation tasks, particularly preparation, analysis, and drafting.

- They gain speed, clarity, and cognitive support, but remain vigilant about bias, confidentiality risks, and loss of human judgment.

- Organisational readiness, training, policies, leadership support, is low, creating a gap that respondents urgently want filled.

- While individuals embrace AI pragmatically, they are cautious about AI's formal role in multilateral diplomacy, calling for governance, ethics, and capacity-building before deeper integration occurs.

* * *