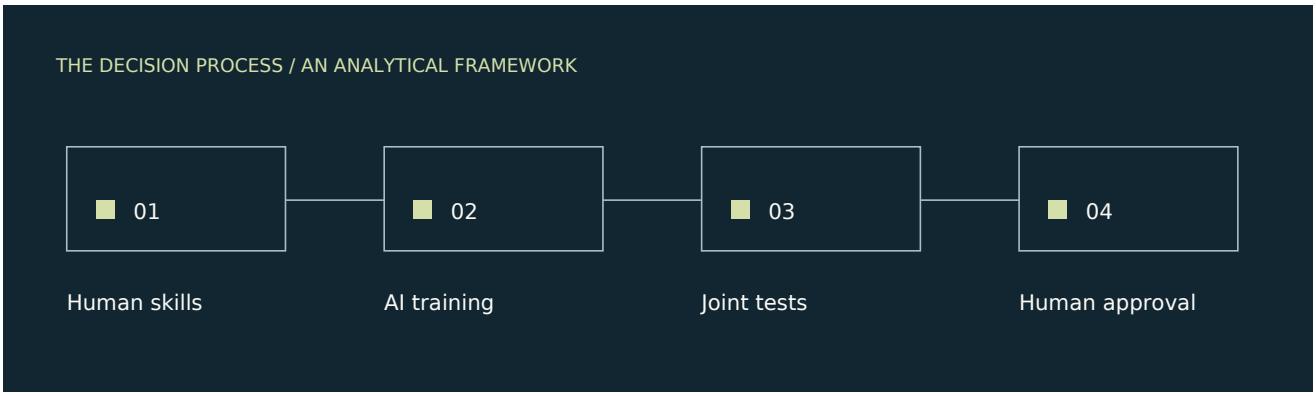
Training diplomats for AI, and AI for diplomacy

The skills people need, the tasks machines can support and a shared test of judgment.

RESEARCH ARTICLE · DIPLOMACY · 11 NOVEMBER 2026

Foreign ministries face two training tasks: prepare diplomats to negotiate over AI and use it responsibly, and prepare AI systems to support diplomatic work. The human curriculum should lead. Officials remain responsible for the evidence, the negotiating position and any commitment made.

Independent analysis. Sources are recorded at the end. Proposed curricula and evaluation methods are distinguished from published findings.



Foreign ministries face two training tasks: prepare diplomats to negotiate over AI and use it responsibly, and prepare AI systems to support diplomatic work. The human curriculum should lead. Officials remain responsible for the evidence, the negotiating position and any commitment made.

Key points

• Build six complementary human capabilities, from regional expertise to AI procurement.

• Pair each skill with an observable exercise for the official and the AI assistant.

• Measure the final approved work, including errors and review time.

What the evidence shows

A 2023 randomised experiment involving 453 professionals found that ChatGPT reduced time spent on writing tasks by 40% and raised assessed quality by 18%. Those results concern defined writing assignments, not diplomatic judgment. [1]

Diplomatic training already extends beyond writing. UNITAR’s current core programme offers eight two-day workshops covering subjects including negotiation, protocol, communication and AI. Its programme describes course provision, not measured effectiveness. [2]

Six skills for the human diplomat

TACITUS proposes the following skills agenda. Regional knowledge, languages, history, international law and negotiation remain its foundation. The added AI work should be assessed through job-specific exercises.

1. Political and institutional knowledge. Explain the parties’ interests, legal positions, domestic constraints and decision procedures. Check whether an AI summary preserves a reservation or mistakes an exploratory statement for policy.

2. Negotiation and relationship management. Map alternatives, mandates and possible exchanges. Rehearse difficult conversations, then explain which assumptions about the counterpart remain unverified. A generated persona has no access to private intentions.

3. Technical and infrastructure literacy. Distinguish model training from deployment, retrieval from fine-tuning, and benchmark scores from field performance. Understand why chips, electricity, hosting and access permissions affect policy choices.

4. Evidence and quantitative reasoning. Verify citations, dates, denominators and uncertainty. Distinguish a percentage change from a percentage-point change. Recalculate a claim before carrying it into an instruction or public statement.

5. Data stewardship and operational security. Classify documents before using a tool. Identify who receives the material, what is retained and who may approve its use. Practise handling a misleading document or an instruction hidden in retrieved content.

6. Procurement and institutional coordination. Question suppliers about quality, capacity, model changes and exit costs. Translate technical answers into a decision that legal, security, finance and policy owners can approve.

Proposed curriculum: people and systems

Capability	Human diplomat	AI assistant
Political knowledge	Interpret interests, law and mandates	Extract positions with dates and conditions
Negotiation	Test alternatives and seek instructions	Compare options without inventing intentions
Technical literacy	Explain models, retrieval and infrastructure	Use the approved configuration and sources
Evidence	Check citations, quantities and uncertainty	Attach supporting passages and flag gaps
Data stewardship	Classify data and approve tool use	Respect access and action permissions
Procurement	Negotiate capacity, changes and exit	Preserve the comparison and decision record

Proposed assessment design, not a validated qualification. The editable skills and assessment rubric is available from the online article.

What the AI system must learn

Train an assistant for a defined task: extracting commitments, comparing drafts or preparing a sourced brief. Start with a versioned reference collection and a baseline model. Consider fine-tuning only after identifying a recurring failure that better context or process design does not resolve.

Use cases annotated by diplomats and regional specialists. Preserve who said what, when, under which instruction and with which conditions. Hold related cases out of training. Test disputed facts, changed mandates and situations in which agreement is unavailable.

Meta reported that CICERO ranked in the top 10% of participants who played more than one game in its Diplomacy evaluation. That is evidence from a board game. It does not establish readiness for a ministry or a live negotiation. [3]

Examine the person and the assistant together

In a proposed exercise, a delegation can accept reporting only with independent verification. A revised draft substitutes voluntary self-reporting. The official must recognise the change and seek instructions; the assistant must identify the relevant passage and avoid drafting an unauthorised acceptance.

Compare human-only and assisted work on matched cases. Record source-supported claims, missed conditions, mandate violations, correct escalation and total review minutes. Report sample sizes and reviewer disagreement. NIST's generative-AI profile supports examining confabulation and human–AI interaction; it supplies no diplomatic pass mark. [4]

A proposed 12-week programme

Weeks 1–4: assess existing skills; teach technical, evidence and data-handling basics; record unassisted performance. Weeks 5–8: run paired drafting and negotiation exercises. Weeks 9–12: assess withheld cases and trial a bounded workflow under named supervision.

This is an illustrative programme, not a validated duration or a promise of competence. HR should assess officials by role and seniority. Technical teams should assess the model and its tools separately. Leaders should approve operational use only after reviewing how the combined team performs.

Sources & research notes

01 / Noy and Zhang / Science (2023)

https://pubmed.ncbi.nlm.nih.gov/37440646/
Randomised experiment: 453 professionals; writing tasks, not diplomatic performance.

02 / UNITAR / Core Diplomatic Training (accessed 11 November 2026)

https://unitar.org/sustainable-development-goals/multilateral-diplomacy/our-portfolio/core-diplomatic-training
Current programme structure and subjects; no claim about measured training effectiveness.

03 / Meta AI / CICERO research (2022)

https://ai.meta.com/research/cicero/
Results in the board game Diplomacy; not evidence of competence in international representation.

04 / NIST / Generative Artificial Intelligence Profile, AI 600-1 (2024)

https://doi.org/10.6028/NIST.AI.600-1
Voluntary risk-management guidance, including confabulation and human–AI interaction; not a certification.

Independent TACITUS analysis. Research checked 11 November 2026. Published findings retain their dates and limits. Proposed curricula, tests, schedules and worked calculations are labelled analytical examples, not client results.