

Recommendations for Managing the Trade-offs of AI in Diplomacy

Over the course of two convenings in 2023 and 2024, a panel of experts from diplomatic organizations, academia, and civil society identified ethical trade-offs implicated in the use of large language models (LLMs) by diplomatic entities and developed a series of proposals to manage these trade-offs. Those proposals are summarized below.

A loss of transparency, traceability, and accountability

It's difficult to trace the cause of—and hold the right people accountable for—harms that might arise from LLM system outputs.

Improve information-sharing requirements among vendors

Especially for high-impact use cases, developing more explicit and stringent procurement standards can help mitigate transparency and accountability concerns.

Only adopt systems for which there is at least an after-the-fact ability to trace back to failure

If no responsibility for a system's failure can be assigned, no party can be held accountable.

Consider changing access to the file

Granting counterparts or constituents a view of the kinds of data being used in these systems can improve transparency.

A rise in inequality

Making the most out of AI will require tailored LLMs and/or extensive skilled staffing. This could put such technologies beyond the reach of less resourced diplomatic parties.

Ensure AI-based initiatives are set up and driven by local groups

Without local support and know-how, AI projects may get off the ground but will not last over the long term.

Improve global understanding of cyber risk management and the global threat environment

If these tools are adopted without cyber risk fluency, security breaches may result in widening inequality gaps.

Train and build AI capacities through international cooperation

Cross-cultural knowledge sharing strengthens multilateralism and protects against rising technical inequalities.

A skill, faded

An organization that leans heavily on AI may gain efficiencies in the short term, but they might lose certain key skills among their staff.

Clearly delineate between tasks where AI should and should not be used

By identifying the tasks where AI should not be used—whether for security, accuracy, or deskilling concerns—safeguards can be put in place to protect against skill loss.

Allocate resources for training the next generation of diplomats

Educating, training, and preparing future diplomats for their tech-enabled future can help keep humans in the loop while reaping the benefits of the technology.

AI Task Suitability Framework

[Download the framework](#) to evaluate the suitability of LLMs for various tasks given your organization's internal priorities, resources, and values.