Computational Statecraft: An Overview

PROJECT ON COMPUTATIONAL STATECRAFT · BELFER CENTER, HARVARD KENNEDY SCHOOL

ABOUT Questions of knowledge, trust, and proof have always been at the heart of the practices of statecraft. Increasing computational capabilities change all of these. Memos, treaties, briefs, and monitoring practices will become intertwined with machine intelligences.

Governments will deploy AI in the practice of statecraft, but it must be ethical, safe, and effective.

01 WHY NOW?

Adoption is running ahead of evaluation. Foreign ministries and defense agencies are already deploying AI tools, whether StateChat in the State Department or CamoGPT for the U.S. Army. AI tools are becoming informally embedded in analytic work, raising concerns about security, over-reliance, and a loss of human judgment.

Nobody is checking claims about AI capabilities in this domain. There are no shared standards for judging whether these tools actually help, and incentives aren't aligned with honesty. AI companies competing for government contracts and the governments buying from them both have reasons not to publicize weaknesses.

If used recklessly, AI could feed unchecked analysis into high-stakes decisions faster than human teams can catch errors. Used well, it could reduce misunderstanding between states and level the playing field between large and small delegations.

02 WHAT WE MEAN BY COMPUTATIONAL STATECRAFT

Computation is becoming the medium of statecraft, not just a tool for it. For decades, computers supported diplomacy in a similar way to telephones: carrying information between people who did the thinking. That has changed. When machines help write the assessment, generate the options, and monitor the agreement, they shape what governments know, and, therefore, what they do. Statecraft is starting to run through computation.

This wave differs from the telegraph or the internet in three ways. Firstly, it touches every domain of statecraft simultaneously — trade, security, energy, negotiation. Secondly, the most capable systems are held by a few companies and countries, creating new dependencies. Finally, for the first time, the technology takes part in the thinking itself, rather than just moving information.

Not everything about world politics can be computed and verified. Machines excel where a task can be clearly stated and its output checked: research, precedent, drafting, monitoring. Humans are essential for interpersonal judgment, building trust, and deciding what counts as success.

The line between work that can be done by a humans and machines is not fixed. Shifting that boundary outward for oneself and inward is itself becoming part of international competition. It may be that advantages will go not to the best AI system, but more to the states that adapt their work (both human and artifical) most wisely.

We call this edge the **verifiable frontier** and advocate that AI systems should be integrated where it can be monitored and verified. Human decisions should remain decisive where needed. Responsibly improving the capacities of human-AI teams is the program's central technical project.

03 THE PROGRAM

We are an interdisciplinary initiative bringing together diplomatic practitioners, social scientists, and technologists. We work across three key tracks:

CONCEPTUAL FRAMEWORKS — a conceptual whitepaper and working papers, including studies of the foreign policy evaluation gap and how practitioners already use these tools.

TECHNICAL RESEARCH — a benchmark suite that tests AI systems against real diplomatic tasks instead of stylized scenarios, so claims about usefulness can be checked.

INSTITUTIONAL GUIDELINES — a primer series for practitioners and an annual State of AI in Diplomacy report, giving policymakers the technical literacy to judge what they are buying.

CONVENING — working groups and policy convenings across the U.S., Europe, and globally, connecting the people who build, deploy, and govern these systems.