

Summary of *Decoding Intentions: Artificial Intelligence and Costly Signals*

In this brief, we explore a crucial artificial intelligence (AI) policy lever that has not received much attention in the public debate: **costly signals**. Policymakers can send credible signals of their intent by making pledges or committing to undertaking certain actions for which they will pay a price—political, reputational, or monetary—if they back down or fail to make good on their initial promise or threat. Talk is cheap, but inadvertent escalation is costly to all sides. By expanding the AI toolkit to include costly signals, policymakers can prevent misperceptions and better communicate intent amid intensifying geopolitical competition.

Costly signals offer a way to communicate intentions in situations of low trust, but they operate differently today than during the Cold War. Based on a close examination of major power signaling on military AI and autonomous weapons, U.S. government signaling on democratic AI, and private sector signaling around the release of powerful language models, this study highlights the following policy considerations and lessons learned.

- *Not all signals are intentional, and commercial actors may conceptualize the costs differently from governments or from industry players in other countries.*
- *Costly signals are one tool in the AI policy toolkit and must be embedded in comprehensive strategies.*
- *The policy choice is not simply whether to “conceal or reveal” AI capabilities, but also how to reveal them and through which channels.*
- *The coupling of public and private sector messaging and actions can be a powerful source of costly signaling.*
- *As a tool of technology policy, costly signals come with trade-offs that need to be managed, including tensions between transparency, privacy, and security.*
- *The ability to convey costly, credible, and clear signals varies depending on the context and technology area.*

Policymakers should incorporate costly signals into scenario planning and table-top exercises with allies and competitor nations to mitigate the risks of inadvertent escalation and develop shared understandings around crisis communication. Signals can be noisy, occasionally confusing some audiences, but they are still necessary.

For more information:

- Download the report: <https://cset.georgetown.edu/publication/decoding-intentions>
- Contact: Andrew Imbrie (wai2@georgetown.edu)