



BGF WEEKLY

Boston Global Forum · AIWS · Week of June 29 – July 5, 2026

SECTION 2 · AI POLICIES AND STRATEGIES

Five Eyes and a New Governance Standard for Agentic AI



"Careful Adoption of Agentic AI Services" – Five Eyes cybersecurity agencies, May 1, 2026

The rapid emergence of agentic AI marks a new chapter in artificial intelligence. Unlike earlier systems that primarily generated content or responded to prompts, agentic AI can independently

plan, make decisions, coordinate with other agents, use digital tools, and execute complex tasks with limited human intervention.

For governments, this presents both extraordinary opportunity and unprecedented governance challenges. On May 1, 2026, the cybersecurity agencies of the Five Eyes nations — the United States, the United Kingdom, Canada, Australia, and New Zealand — took a first coordinated step, jointly publishing “Careful Adoption of Agentic AI Services,” the first multigovernment security guidance devoted specifically to autonomous AI systems. It identifies five categories of risk — privilege, design and configuration, behavioral, structural, and accountability — and insists that strong governance, explicit accountability, rigorous monitoring, and human oversight are essential prerequisites, not optional safeguards.

The Boston Global Forum welcomes this milestone — and believes the moment calls for more. While existing frameworks emphasize safety, transparency, privacy, and accountability, systems capable of autonomous action must do more than operate safely; they must remain trustworthy throughout their operational lifecycle. The next generation of AI governance should move beyond AI safety toward a new international standard built on Trust Infrastructure.

Such a standard should rest on several foundational principles:

- ◆ **Human Authority.** Every agentic AI system must operate under clearly defined human authority, with meaningful human oversight of critical decisions.
- ◆ **Trust by Design.** Trust should be engineered into the architecture of AI systems from the outset, rather than added after deployment.
- ◆ **Continuous Trust Evaluation.** Agents should be continuously assessed through measurable trust indicators — reliability, factual integrity, cybersecurity resilience, behavioral consistency, and compliance with established governance standards.
- ◆ **Agent Accountability.** Every autonomous action should be traceable, explainable, and auditable through secure logging and provenance mechanisms.
- ◆ **Collective Governance.** Multi-agent ecosystems require governing not only individual agents but also their interactions, coordination, and emergent behaviors.

These principles align closely with the AIWS Trust Infrastructure. Rather than relying on regulation alone, AIWS proposes a practical governance architecture combining Trust Ratings, Trust Standards, Information Trust Infrastructure, continuous monitoring, and human-centered oversight.

As agentic AI becomes an essential component of government, business, healthcare, education, finance, and national security, the international community faces a choice: it can continue extending frameworks designed for earlier generations of AI, or it can establish a new paradigm built specifically for autonomous intelligent systems. The emergence of agentic AI calls not simply for stronger regulation, but for a new architecture of trust.

In the AI Age, trust is no longer merely an ethical aspiration. It is becoming essential public infrastructure.

The Five Eyes guidance is an important opening. It should evolve beyond national-security cooperation toward internationally interoperable governance standards that democratic societies can share with partners around the world. The Boston Global Forum will continue advancing AIWS Trust Infrastructure and working with governments, technology leaders, and international organizations to help build this new foundation for trustworthy agentic AI.

BGF Lens — *the Five Eyes guidance treats agentic AI as a security problem; the Boston Global Forum reframes it as a trust problem. Security keeps autonomous systems from being exploited; trust ensures they remain worthy of the authority we grant them — with the human person always in command.*

Sources: “Careful Adoption of Agentic AI Services,” Five Eyes cybersecurity agencies, May 1, 2026; Boston Global Forum · AIWS Trust Infrastructure.