# THE CONVERGENCE

## FROM DIFFERENTIATOR TO GATE: FORMAL METHODS AND THE 2027 AI PROCUREMENT SHIFT

The technical and legislative convergence that led to formal verification becoming the price of entry to AI procurement in the US Government.

**July 2026**

**KRY10 WHITE PAPER**

*By Jason Fox, Co-Founder + COO*

kry10.com
USA | UK | AUS | NZ

## Executive Summary

**Formal methods, the discipline of mathematically proving a system does only what it is designed to do, has spent seventeen years moving from an academic proof of a single kernel to a defense-wide requirement. In December 2025, the FY2026 National Defense Authorization Act (P.L. 119-60) wrote it into law: a DoD-wide governance policy for AI/ML systems, procurement security requirements, an AI assessment framework, and an AI sandbox task force. DARPA's own research funding solicitations now score proposals on whether they use formal methods.**

**At the same time, AI capability has spent fourteen years outrunning the government's ability to govern it by agreement alone. That gap became quickly evident in June 2026, when the US government banned Anthropic from operating its most capable models over an exploit nobody had tested for. Access was restored, but only partially, and only under an invitation-only regime. The underlying problem, that capability keeps outrunning proof, was not solved.**

**These two timelines are converging. The US government is moving, through binding acquisition law, toward requiring mathematical proof rather than confidence for the AI systems it buys and fields. Formal methods are becoming the answer to the AI problem, and the organizations that already have a formally verified foundation, rather than one they are hoping to retrofit, will be the only ones ready when that requirement lands.**

## Introduction

Every AI safety debate eventually hits the same question:

**How do you actually know a system won't do the catastrophic thing?**

Red-teaming and testing both find some failures, but neither can prove the absence of the failure that nobody thought to test for. That gap, between "we tried hard to break it" and "we can show mathematically that it can't do that," is driving the current tension between Washington and frontier AI labs, most visibly in the recent export-control standoff over Anthropic's Mythos and Fable models.

The US Government is effectively asking, "is this safe to deploy?" The industry currently doesn't have a rigorous way to answer that, but formal methods does. This White Paper explores the story of two timelines, one technical, one political, that have spent seventeen years heading toward exactly this moment.

## Formal Methods Background

**Formal methods (mathematically proving a system behaves as specified under all conditions, not just the ones you happened to test) has been building toward large scale adoption for seventeen years.**

### A short history of formal methods

- In 2009, researchers completed the first machine-checked proof of functional correctness for a general-purpose kernel, seL4® demonstrating that the approach could scale past small, focussed problems.
- Aviation certification was next, when DO-333 (2011) made formal methods an accepted path to certifying safety-critical avionics code.
- DARPA proved it works under attack with its High-Assurance Cyber Military Systems (HACMS) program (2012–2017), producing a Boeing helicopter that a red team, with far more access than a real attacker would ever get, couldn't hack past its perimeter.
- By 2024, the White House's Office of the National Cyber Director (ONCD) named formal methods a national building block for eliminating entire vulnerability classes. And DARPA's Pipelined Reasoning Of Verifiers Enabling Robust Systems (PROVERS) program, running now, is specifically about getting formal methods out of specialist hands and into ordinary engineering teams.

By 2026, the use of formal methods started to accelerate.

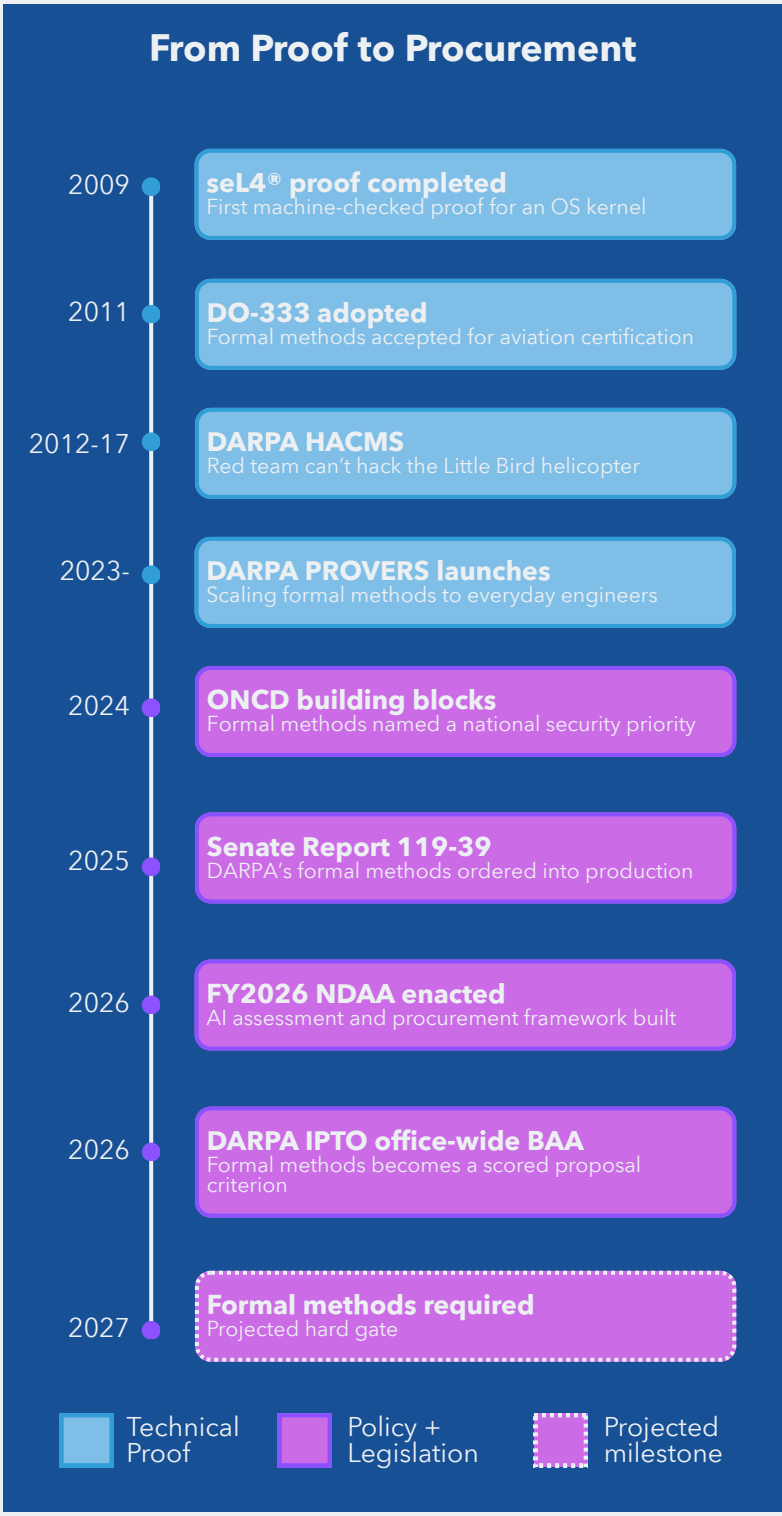


**Figure 1: Formal Methods Timeline**

## AI Capability Growth

**AI capability, meanwhile, has been outrunning anyone's ability to test it for almost as long, and so far every government response has been focussed around conversation, not action.**

### The exponential growth of AI capability

- In 2012, AlexNet's win at ImageNet proved that deep learning could scale past academic toy problems This was the same year that DARPA started HACMS on the formal methods side.
- In 2017, "Attention Is All You Need" introduced the Transformer, the architecture every frontier model since has been built on.
- In 2022, ChatGPT went from launch to a million users in five days, taking frontier AI to the masses faster than any prior technology.
- Between 2023 and 2025, the Bletchley Summit, Executive Order 14110, the EU AI Act, and America's AI Action Plan each attempted to govern frontier AI by agreement and commitment, not by proof.

Then in June 2026 the gap between AI capability and our ability to govern and test became a crisis. The US Government banned Anthropic from operating Fable 5 and Mythos 5 over an exploit nobody had tested for.

Fourteen years of capability outrunning governance had produced summits, executive orders, and action plans. But the US Government's decision in June 2026 produced an actual takedown, and even at the time of writing, the most capable model stays behind an invitation list, not a public release.

Gating one company's model doesn't gate the capability. Anthropic itself has said that the same exploit technique worked against other publicly available models, including OpenAI's GPT-5.5, which faced no restriction at all.

The problem is coming, and it's now time to decide how to deal with it.


**Figure 2: AI Timeline**

## THE FMxAI CONVERGENCE

**Two timelines are colliding here.**

Formal methods moving from a single proven kernel to a DARPA program built to put mathematical proof in the hands of ordinary engineering teams, not just specialists, settles a narrow technical question about whether you can prove a complex system does only what it's supposed to do. The answer is now yes, at scale.

AI systems are shipping faster than anyone can test them, but testing is never going to be the thing that catches the failure nobody thought to check for. That was the tension between Washington and Anthropic over Mythos and Fable. The US Government asked "is this safe to deploy?" and industry was unable to answer with anything stronger than "we tried hard to break it."

Formal methods and AI are converging because they're two halves of the same unsolved government problem: **how do you approve a system for deployment when testing can tell you it failed but can't tell you that failure is impossible?**

Formal methods is the one discipline built to answer with proof instead of confidence. That's why, when we look at our timelines, formal methods stopped being a specialist cybersecurity niche around the same time that frontier AI stopped being an experiment.

One of the next big moves for formal methods won't be happening in a lab, it will be a change in procurement legislation.

## National Defense Authorization Act (NDAA)

In the United States, that acquisition law arrived in December 2025, when the The National Defense Authorization Act (NDAA) was signed into law as P.L. 119-60. The NDAA is the annual law that sets policy and authorizes funding for the Pentagon. Anything directed in the NDAA or its accompanying committee report isn't a suggestion, it's tied to the mechanism that funds and justifies Department of Defense (DOD) programs every year.

That's what makes Senate Report 119-39 (the Armed Services Committee's report accompanying the FY2026 NDAA, July 2025) so significant. This report directed DOD to develop "a comprehensive strategy for transitioning DARPA's formal methods research investments into production environments across the DOD."

In effect, this means that current usable formal methods work have now been ordered into programs of record, with the NDAA itself supplying the mechanics to enable it.

## Where the convergence starts to impact

The specific areas where the NDAA sparks the FMxAI convergence are Subtitle B, Sections 1512 and 1513 and Subtitle D, Sections 1533 and 1534.

In essence Section 1512 requires DOD to build a governance policy addressing model tampering, data poisoning, unauthorized access, and supply chain attacks against AI/ML systems. Meanwhile, Section 1513 requires physical and cybersecurity procurement requirements for those same systems. Both sections are targeting assurance that a specific AI system running on specific hardware hasn't been tampered with and can't be manipulated. That's precisely the problem a formally verified, isolation-enforcing microkernel and operating system (OS) answers at the platform layer. It ensures that tampering and unauthorized access are mathematically impossible to achieve through a formally verified kernel (like seL4®) and OS (like KOS, Kry10's operating system), not by bolting security onto an AI stack after the fact.

Section 1533 is where the actual DOD-wide AI assessment framework gets built, including the testing procedures, security requirements, and the mechanism that would eventually score or gate based on formal methods evidence. This is the section that determines what "proof" looks like in practice, and the standard AI solutions would need to be evaluated against. Section 1534, the AI sandbox task force, is a more specific opportunity: an seL4®/KOS-based isolation is well suited to building strongly partitioned sandbox environments for AI experimentation and implementation.

It's important to read the NDAA and Senate Report 119-39 together. The NDAA is the mechanism, the Senate Report 119-39 is the directive that names formal methods, while Subtitle B, Sections 1512, 1513 and Subtitle D, 1533, and 1534 are where that directive starts to impact. Subtitle B defines the problem formal methods solves, proving an AI system hasn't been tampered with (1512) and turning that into procurement requirements (1513). Subtitle D defines where the proof gets tested, the assessment framework AI systems will eventually be scored against (1533), and where formally verified systems get a first venue to prove themselves (1534). Read against Senate Report 119-39, these sections are the acquisition machinery already built to receive formal methods.

## Formal methods and AI today

In June 2026, DARPA's Information Processing Techniques Office (IPTO) published a live, office-wide solicitation (Broad Agency Announcement, or BAA, HR001126S0011, June 22, 2026) stating plainly that "DARPA's preferred method" for delivering resilient software "is the use of formal methods," and making it an actual scored criterion. Proposals are evaluated in part on whether they use formal methods to eliminate vulnerabilities. The directive and the practice are arriving at the same time.

The BAA itself is still DARPA picking who gets research funding, a factor that helps a proposal's score, but is not yet a gate on whether a fielded system gets bought. Sections 1533 and 1513 are aimed at production acquisition directly, not research grants. A DOD-wide AI assessment framework and AI/ML procurement security requirements apply to programs of record - the systems the Pentagon actually buys and fields. Both tracks, R&D funding and production acquisition, are already moving. Research funding and production acquisition for AI/FM are moving at the same time. That's the convergence.

## Formal methods and AI in 2027

What tomorrow is expected to bring is the shift from optional-and-scored (research) to required-and-gating (acquisition). Formal methods stops being a factor that helps a proposal win and becomes a condition, whereby nothing gets purchased by the US Government without it.

Industry conversations are pointing to 2027 as the year this shifts from being a differentiator to a prerequisite, but the direction is clear. Formal methods is on an irreversible track toward becoming a legislated requirement of doing business with the US Government on anything AI-adjacent. Once that lands, every Tier One Defense Company, serious CNI owner and DOD program running AI at the edge will need a foundation that's already proven, not one it's hoping to retrofit proofs onto after the fact.

The Kry10 OS, built on the seL4® microkernel, is that foundation today. While most of the industry will spend 2027 scrambling to bolt assurance onto AI stacks that were never built for it and AI/FM companies sprout up claiming they can produce miracles in a few months, they will all lack one important thing. You have to start from a foundational proof that is usable, and this is where Kry10's KOS is the ideal secure substrate for  AI and everything else to sit on.

Confidence was always going to run out eventually. Proof is the only thing that scales past it.

## About Kry10

At Kry10, our job is to strengthen the foundations beneath the critical systems that society depends on every day. Kry10 provides the mathematically proven base to help keep critical services resilient by design.

KOS, our formally verified configuration operating system, is built on the formally verified seL4®, the world's most highly assured operating system kernel.

Kry10 works with government and private organizations charged with the cyber defense of critical assets. The Kry10 team brings together world-leading expertise at the intersection of formally verified operating system design, microkernel architecture, and OT cyber security.

The result for our clients is the freedom to focus on what matters most: serving the world long after today's threats have passed.

## Sources

U.S. Senate Committee on Armed Services, National Defense Authorization Act for Fiscal Year 2026, Report 119-39, July 15, 2025, p. 310;
FY2026 NDAA (P.L. 119-60), Sections 1512-1513, 1533; CRS In Focus IF13197; seL4 project history; RTCA DO-333 (2011);
DARPA HACMS case study; ONCD "Back to the Building Blocks" report, Feb. 2024; DARPA PROVERS program;
DARPA IPTO Office-Wide BAA HR001126S0011, June 22, 2026; DARPA, "Ushering in the new era of cyber resilience," June 26, 2025;
Krizhevsky, Sutskever & Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NeurIPS 2012;
Vaswani et al., "Attention Is All You Need," 2017; OpenAI, "Introducing ChatGPT," Nov. 30, 2022;
UK Government, "The Bletchley Declaration," Nov. 1, 2023;
Executive Order 14110, "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," Oct. 30, 2023; Regulation (EU) 2024/1689 (EU AI Act);
The White House, "Winning the Race: America's AI Action Plan," July 2025;
Anthropic, "Statement on the US government directive to suspend access to Fable 5 and Mythos 5," June 12, 2026;
Jeremy Kahn, "Anthropic disables Fable and Mythos AI models after U.S. government bars it from giving foreigners access," Fortune, June 13, 2026;
Anthropic Confirms Claude Mythos 5 Redeployment for US Critical Infrastructure Organizations, June 2026.

**We're ready to partner**

**Let us show you what verifiable resilience looks like in practice**

hello@kry10.com
kry10.com
USA | UK | AUS | NZ