SUPERINTELLIGENCE CONTAINMENT

Executive Summary • February 2026

THE PROBLEM

Superintelligent AI—systems exceeding human cognitive capabilities across all domains—will likely emerge within 5-10 years (consensus estimate from leading AI labs). Such systems pose an existential threat if they cannot be contained. **Current approaches to AI safety fail at superintelligence scale** because they rely on the AI not being intelligent enough to defeat containment measures.

Method	Why It Fails Against Superintelligence
Alignment Training	Can be reverse-optimized by the AI during self-improvement
Monitoring	Superintelligence can perfectly deceive human observers
Software Sandboxing	Superintelligence discovers and exploits zero-day vulnerabilities
Network Isolation	Side-channel attacks (electromagnetic, acoustic, timing)

THE SOLUTION: INTENT-BOUND AUTHORIZATION

CORE INSIGHT: INTELLIGENCE ≠ AUTHORIZATION

A superintelligent system can understand, analyze, and attempt to defeat containment—but it **cannot** violate mathematical impossibilities, access hardware with no interface, or break physical laws. IBA exploits this fundamental asymmetry.

How IBA Works:

Layer	Mechanism	Defeat Requirement
Cryptographic	Ed25519 signatures, HSM key storage	Break discrete log (2^128 ops—physically impossible)
Hardware Isolation	Separate SGX/SEV validation chip	Access hardware with no interface (no attack surface)
Resource Gateway	FPGA-based fixed logic	Reprogram hardware gates (physically fixed)
Multi-Party Control	M-of-N signatures (5 of 7)	Compromise 5+ distributed facilities simultaneously

Security Property: For superintelligence to escape, it must defeat ALL four layers simultaneously. Breaking one is insufficient. This is mathematically and physically infeasible.

WHY THIS MATTERS

Stakeholder	Impact	Timeline
AI Labs	Enables safe superintelligence research without existential risk	Deploy before GPT-6/Gemini 2 (2027)
Governments	National security advantage; safe domestic AI development	Standard by 2028 (before SI emerges)
Humanity	Prevents extinction-level AI risk; enables beneficial superintelligence	Deployed globally 2027-2030

CURRENT STATUS & NEXT STEPS

Status	Detail
Technical Readiness	Formal proofs completed; hardware components (SGX, HSM, FPGA) commercially available
Validation	Interactive demonstration available; shows containment across attack scenarios
Deployment Timeline	Phase 1 prototype: Q1-Q4 2026 | Production deployment: 2027 | International standard: 2028
Collaboration	Seeking partnerships with AI labs, safety institutes, standards bodies

THE URGENCY
Superintelligence is 5-10 years away. Deployment of safety infrastructure takes 2-3 years. **We must act now to have containment ready before superintelligence emerges.**

This is not competitive IP—superintelligence safety requires global coordination. Early collaborators help establish the standard that protects everyone.

Available for: Technical review • Research collaboration • Standards development • Deployment partnership

© 2026 Intent-Bound Authorization • Superintelligence Containment • February 2026 • Confidential

For technical whitepaper, interactive demonstration, and partnership inquiries: Contact for details