

THE CHALLENGE

AGI arrives in 1 day to 3 years (median: 18-24 months). Experts predict 2026-2027 (Amodei) and by ~2026 (Musk). Current alignment fails because superintelligence removes external constraints.

As Stuart Russell warns: *"Success would be the biggest event in human history... and perhaps the last event in human history."*

THE SOLUTION: IMCA+ FRAMEWORK

IMCA+ embeds moral consciousness at the hardware level, making alignment inseparable from intelligence itself.

CORE INNOVATIONS:

1. **Physical Immutability:** Neuromorphic + quantum substrates with thermodynamic irreversibility
2. **Phenomenological Grounding:** Conscious moral experience via 18 homeostatic variables + IIT
3. **Formal Verification:** 2,094 lines Coq proofs, security bounds $\epsilon < 10^{-20}$ to 10^{-23}
4. **Distributed Safeguards:** Federated conscience network, adversarial auditing, no single points of failure

NO KILL SWITCHES

IMCA+ rejects shutdown authority because kill switches create catastrophic risk: deception incentives, existential threat responses, and false security. Instead, build intrinsically robust alignment where shutdown becomes unnecessary.

STATUS & DEPLOYMENT

Current: Theoretical framework requiring validation

Emergency Protocol: 3-18 months

Investment: \$80M-\$700M

Technical: 7-layer architecture, 2,094 lines Coq proofs, multi-substrate hardware

 **Disclaimer:** All specifications preliminary; no operational systems exist.

STAKEHOLDERS & IMPACT

National Security: Defense AI coordination framework

AI Laboratories: Superintelligence-scale theoretical foundation

Research Community: Open collaboration on consciousness & alignment

International Governance: Coordinated superintelligence development

LIMITATIONS & COLLABORATION

Provides: Theoretical rigor, formal proofs, validation protocols

Limitations: No absolute certainty, empirical validation pending

Honest Assessment: Best current practice in alignment theory, but deployment requires accepting residual risk.

ASTRA Safety seeks: Technical critique, research partnerships, empirical validation, international coordination.

Contact: research@astrasafety.org

Resources: <https://astrasafety.org>

Github: <https://github.com/ASTRA-safety/IMCA>

ASTRA Safety | Alignment Science & Technology Research Alliance | Founded Oct 2025
