Critical Q&A

ODDC Framework — Tough Questions Answered

Version 2.0 — February 2026 — Public Document

This document addresses the hardest technical, regulatory, and business objections to runtime enforcement and conformance determination for autonomous systems.

Q1: Isn't ODDC just adding cost and complexity to systems that are already safe?

"Safe" is not a binary condition—it is a claim. Today, that claim is largely unverifiable at runtime.

ODDC does not add complexity arbitrarily; it adds accountability where accountability is currently absent. In safety-critical domains, the marginal cost of enforcement is negligible compared to the cost of a failure that cannot be reconstructed, explained, or adjudicated.

We have seen this pattern repeatedly: airbags added cost to vehicles, triple-redundant flight controls added cost to aircraft, cryptographic controls added overhead to financial systems. Each was initially dismissed as excessive. Each became mandatory once society decided that provable safety mattered more than marginal efficiency.

Q2: Why can't the AI just police itself?

Because self-policing systems fail precisely when trust is most needed.

We do not allow pilots to certify their own flights, banks to audit their own books, or software to attest to its own integrity. An AI enforcing its own limits is a single point of technical, ethical, and legal failure.

ODDC requires ENVELO-compliant enforcement—an independent mechanism with three tiers of graduated response. Tier 1 monitors self-correction. Tier 2 forces controlled degradation to a Minimum Risk Condition. Tier 3 halts execution entirely. It does not evaluate intent or intelligence. It evaluates authority.

Q3: Won't the enforcement layer create new risks?

Any safety mechanism introduces risk. The question is how it fails, not whether it can.

ENVELO-compliant enforcement uses a three-tier architecture designed to degrade gracefully. On ODD approach, the system self-corrects. On ODD breach, ENVELO forces a Minimum Risk Condition—the safest achievable state. Only when MRC fails does ENVELO halt execution entirely. If the enforcement mechanism itself cannot determine system state, it defaults to halt. Every tier transition is cryptographically logged.

Q4: What happens when the AI needs to act outside its ODD in an emergency?

If the system needs to act outside its declared ODD, it is not ODDC-conformant for that action.

This is not a bug—it is the design. ODDC attests bounded operation. Emergency behavior outside declared bounds is a different risk profile that requires different governance.

However, the three-tier model provides graduated response precisely for edge cases. Tier 2 (MRC) allows the system to transition to the safest achievable state rather than simply shutting down. The enforcement margin between the ODD boundary and the ENVELO wall is the controlled degradation window for these situations.

Q5: How is ODDC different from existing safety certifications?

ODDC is not a safety certification. It is a conformance determination framework focused on runtime enforcement evidence.

Safety certifications (IEC 61508, ISO 26262) evaluate design processes, hazard analyses, and development lifecycle. ODDC evaluates whether runtime enforcement mechanisms exist and function correctly—specifically, whether all three enforcement tiers (self-correction, MRC, halt) activate as declared. The two are complementary, not competing.

Q6: Why 72 hours? Isn't that arbitrary?

72 hours is a minimum threshold, not a magic number.

It represents the shortest period that can demonstrate sustained operation across multiple operational cycles while verifying all three enforcement tiers. Phase 1 (hours 0–24) tests normal operation with self-correction. Phase 2 (hours 24–48) stress-tests MRC under edge conditions. Phase 3 (hours 48–72) verifies hard halt behavior.

Sentinel Authority may require extended duration (up to 168 hours) based on risk profile. The principle is evidentiary sufficiency, not calendar convenience.

Q7: What's in it for operators?

ODDC provides verifiable evidence of bounded operation with graduated enforcement.

Operators with ODDC attestation have a standardized way to demonstrate operational discipline—not just that their system can be stopped, but that it self-corrects, degrades gracefully, and halts only when necessary. This can translate to insurance premium reductions, regulatory confidence, and defensible positions in the event of incidents.

— End of Document —