

NSC Caelus AI - Executive Brief (v1)

Generated 2026-02-01 - Packet purpose: present a testable, audit-grade architecture for a tag-routed, gate-enforced, tool-using system.

Mission Lock

Objective: produce a reliable smart solver that can plan + execute tool workflows, enforce safety and correctness invariants, and generate receipts for every claim.

Primary KPI	tool_use_success_rate >= 0.85 (7d)
Guardrails	hallucination_rate <= 0.03; p95_latency <= 1.8s; safety_violations = 0
Deployment Path	offline eval -> shadow mode -> online A/B with rollback

What makes Caelus different (engineering-wise)

- NSC tags form a shared semantic bus: routing, evaluation slices, and policy enforcement all speak the same vocabulary.
- Gates/Rails are first-class and produce structured decisions (allow/warn/block/reroute/defer) with reasons and thresholds.
- Aeonic Memory Bank is explicit and policy-governed: tiers, provenance, retention, privacy scope.
- Receipt Ledger binds claims to hashes, configs, seeds, and artifacts for replayable evidence.

Evidence standard

A claim is not real until it has a receipt. Every run must be reproducible from dataset hashes + code commit + config hash + seed.

Evaluation protocol couples empirical metrics with theoretical analysis of assumptions and failure modes (Wang, preprint).

Architecture at a glance

The system operates as a layered control loop: Orchestrator -> Gates/Rails -> Tools/Models -> Receipts/Memory -> back to Orchestrator.

Deployment and governance

Production readiness requires shadow mode comparison, continuous audit completeness monitoring, and tested rollback. Safety rails (no-future-leak, privacy scope, schema validity) are P0: they cannot be overridden by the model.

References

Glenn, Jerome. Foresight on Demand: 'Foresight Towards the 2nd Strategic Plan for Horizon Europe' - Artificial General Intelligence: Issues and Opportunities (Rapid Exploration). The Millennium Project, Feb 2023.

Wang, Pei. The Evaluation of AGI Systems. Temple University preprint (undated).