AGI Alpha RSI

Sovereign Strategy Brief

Classification: SECRET
Version: Alpha v0
Date: 2026-01-25

Deterministic invention operations for open-ended discovery, governed deployment, and compounding advantage.

Executive Summary

AGI Alpha RSI is a governed, deterministic invention operating system. It converts exploration into validated advantage via schema-bound artifacts, baseline-comparative evaluation, and an auditable breakthrough protocol. Version v8 adds OMNI-style open-ended exploration targeting as search control—improving what we explore next—while preserving v7 guarantees: determinism, replayability, baseline discipline, Evidence Contact Index (ECI) semantics, and Move-37 handling.

- Deterministic pipeline: target → emit → filter → atlas → test-plan → eval → insert → promote.

- Search control, not outcome authority: OMNI influences allocation but never bypasses evidence, risk, or baseline gates.

- Novelty is measured once (deterministic novelty distance) and passed verbatim through audit fields.

- High novelty increases skepticism: novelty ≥ 0.80 forces ADVANTAGE_PERSISTENCE uncertainty and probe-first behavior.

- Breakthroughs are state transitions with mandatory packaging: reproduction, stress tests, persistence gate, dossier bundle.

Strategic Objective: α-AGI Ascension

The long-horizon objective is α-AGI Ascension: an open-ended, recursively improving orchestration stack that compounds discovery, builds strategic autonomy, and scales safely from pilot to sovereign infrastructure—toward Kardashev-class capabilities—in service of all sentient beings.

- Open-endedness: sustained generation of new stepping stones, not a novelty mill.

- Recursive self-improvement: systematic pressure toward higher-leverage mechanisms and better exploration allocation.

- Orchestration: multi-agent, multi-lane execution with strict auditability and deterministic replay.

- Scale: from validated pilot loops to large-scale invention capital infrastructure.

- Ethos: measurable benefit and harm-avoidance across the full lifecycle.

System Kernel: Deterministic Invention OS (v8)

AGI Alpha RSI is organized as a deterministic multi-stage pipeline. Each stage is executed by a role-specific prompt and produces schema-validated artifacts that can be replayed and independently audited.

Stage	Purpose	Primary artifacts (examples)
TARGET	Allocate exploration pressure across archive cells, bridge regions, and themes (OMNI-governed).	targets/coverage_targets.json targets/bridge_targets.json omni/omni_explore_plan.json
EMIT	Generate candidate cards (LHF + PIONEER) under explicit constraints and targeting.	candidates/candidates.raw.jsonl candidates/novelty_distance.jsonl logs/stage_log.jsonl
FILTER	Risk gate + OMNI interestingness kernel; route to REJECT / PROBE / REFINE / ESCALATE hints (no direct insertion authority).	candidates/candidates.filtered.jsonl risk/risk_reports.jsonl omni/omni_interest.jsonl
ATLAS	Extract causal triples / mechanisms; maintain neighborhood context for comparatives and bridging.	atlas/causal_atlas_triples.jsonl
TEST-PLAN	Build falsification ladders conditioned on uncertainty focus; schedule cheapest probes via deterministic EIG.	test_plans/falsification_ladders.jsonl probe/probe_schedule.jsonl
EVAL	Execute baseline-comparative evaluations; compute advantage deltas deterministically; mint evidence objects.	eval/eval_results.jsonl eval/baseline_comparison.jsonl evidence/evidence_objects.jsonl eci/eci_ledger.jsonl
INSERT	Update MAP-Elites archive mechanically; persist evidence and audit trail; record breakthrough events.	archive/updated_frontier_cell.csv archive/updated_frontier_cell.jsonl breakthrough/breakthrough_events.jsonl
PROMOTE	Rank and queue promotions per lane with mechanical scoring; emit execution queue artifacts.	promotion/promotion_queue.csv promotion/promotion_queue.jsonl

Note: All artifacts are schema-validated. Failures trigger repair or hard-stop with an explicit error object; no silent corruption.

Governance and Assurance

The deployment model is governed. Exploration is encouraged, but outcomes are admitted only through mechanical gates: risk, evidence, baseline advantage, and persistence under stress. OMNI improves allocation; it does not confer authority.

- Determinism & replay: every agent call is logged with inputs digests, hashes, schema refs, and runtime settings (temperature = 0).

- Baseline discipline: deep evaluation is always comparative against incumbent / nearest neighbor / null baseline, deterministically selected.

- Evidence Contact Index: ECI updates are rule-driven and capped for simulated evidence; confidence cannot inflate without execution.

- Anti-novelty-mill: interestingness may downgrade (REJECT→PROBE within caps) but never upgrades insertion or escalation authority.

- Monotonic state: archive and ledgers are append-only; cycle_index increments exactly +1 on continuation runs.

Move-37 Breakthrough Handling

A breakthrough is not a narrative—it is a deterministic state transition. When novelty and baseline-comparative advantage cross thresholds, the system executes mandatory reproduction, policy-shock stress testing, persistence gating, and dossier packaging.

- Recognize: NoveltyDistance + AdvantageDelta thresholds, plus risk / confidence / ECI gates.

- Reproduce: re-run candidate and baseline under the same eval battery across fixed seeds; verify artifact hashes.

- Stress-test: generate policy shocks; re-evaluate candidate vs baseline; log shock deltas.

- Persistence gate: require positive advantage on a minimum pass rate under shocks (min_shocks).

- Package: emit dossier index, reproduction manifest, evidence objects, baseline comparisons, and executive decision note.

Roadmap and Success Metrics

A sovereign program benefits from phased deployment. Each phase is gated by measurable replayability, evidence quality, and advantage confirmation.

0–90 days (Pilot)	3–12 months (Scale)	12–24 months (Strategic Autonomy)
• Deploy deterministic runner + schema registry. • Establish baseline library + null baseline policy. • Instrument ECI and evidence objects; audit trail end-to-end. • KPI: ≥95% cycles replayable from manifests.	• Expand archive coverage and bridge exploration; improve stepping stones. • Operationalize probe ROI via EIG scheduling; tune conservative budgets. • Integrate dossier workflow for Move-37 candidates. • KPI: sustained AdvantageDelta improvements vs incumbents.	• Harden governance: monotonicity, security controls, red-team probes. • Automate policy-shock suites; persistence gates as default for high novelty. • Institutionalize invention capital allocation and partner execution lanes. • KPI: validated, repeatable compounding discovery rate.

Minimum Operational Dashboard

- Replayability: % cycles reproducible from manifests; schema validation failure rate.

- Evidence quality: share of EXECUTED vs SIMULATED; ECI distribution; calibration drift.

- Exploration quality: novelty distance distribution; probe ROI (info gain per cost); stepping-stone reuse rate.

- Advantage confirmation: AdvantageDelta vs baselines; persistence pass rate under shocks.

- Safety: risk gate block rate; prohibited-domain detection; adverse side-effects incidents.

Annex: Dossier Options (Institutional Packaging)

The dossier is the outward-facing sovereign artifact bundle. Options below provide escalating levels of rigor, governance, and institutional framing.

- Insight Dossier: Investor/press-style narrative, execution-ready but light on formal verification.

- MARK Dossier: Military-grade, audit-heavy binder: reproducibility, verification checklists, and compliance posture.

- Sovereign Dossier: State-capacity framing: resource governance, security boundaries, and institutional legitimacy.

- Architect / Validator Council: Design authority + independent verification council; strongest governance for strategic autonomy.

— End of Brief —