

AGI-COS Controlled Version Family

AGI-COS | Controlled Version Family | One-page staged entry V1.0

Choose the real-world problem, stage, phase, and task first, then the version. Do not choose a text by the highest version number.

Core judgment: AGI-COS has no single “highest” or “latest” version covering every text. Its current structure consists of cross-version invariants, specialist and research tracks, accompanying constraints and formulations, and derivatives for real-world tasks. Tracks can be used in combination, but combining them cannot accumulate permissions. ACTIVE does not mean automatically enabled.

What are you trying to solve?

Real-world problem or stage	Preferred entry	Main use	Do not inherit automatically
All high-consequence judgments	T0 + 1.1	Cross-version safeguards; sources, confidence, unknowns, and reasons for low confidence. First reduce model outputs to auditable advice.	Any direct conversion of a risk score or model conclusion into a decision or permission.
Crisis or high irreversibility	Add 1.0 + 1.2	Detect → Diagnose → Window → Impact → Path → Execute → Recover → Learn; adversarial data, cognitive load, and second-order feedback.	Historical fixed thresholds such as 48 hours or effects on one million people; psychological persuasion or behavioral shaping.
Present or near-term coordination across actors	2.0 + 2.1 + 1.1	Continuous risk sensing, recovery and learning, distributed trials, non-sovereign coordination, shared risk awareness, and reversible responses.	Fixed years, seats, or participation thresholds; treating MEU as validation of an entire civilization; covert micro-interventions.
Real deployment or interface prototypes	2.2 + 1.1 + current governance	Candidate interfaces for identity, authorization, auditing, sunset provisions, validation nodes, fail-safe behavior, privacy, and related functions.	Generalizing historical implementations such as <100 ms, a 72-hour default, a particular blockchain, IPFS, or a Dead Man's Switch.
Stable conditions or post-AGI long-term design	3.0 + 2.0 + LC-1.0	Subjecthood, the right to fail, an open future, civilizational memory, and systems stepping back; less explicit intervention after stability is achieved.	Exact civilizational utility formulas, meaning scores, or fixed global activation.
Resilience or the culture of failure	3.1 research track	Failure records, Shadow Audit, and resilience with low technology or without network access.	Public Entropy Log, SII/NCS governance scores, or mandatory training or qualifications.
Once real conditions for interstellar activity exist	4.0 / 4.1 research tracks	Distributed coordination, protocol compatibility, memory, drift, and repair over long distances.	Importing multiplanetary assumptions into the present; allowing AGI ethics to override sovereignty; adjusting governance powers by entropy or contribution.
Very-long-term philosophical stress testing	Ω / $\Omega+$	No final endpoint, regeneration of meaning, openness to the unknown, guardian recursion, and resistance to rigidity.	A sole purpose of continued existence, Challenge Injection, manufactured crises, or mandatory Goal Obsolescence.

Seven questions that must be answered before every use

- 1 What is the real-world problem: crisis, routine risk, coordination, deployment, the long term, interstellar activity, or the very long term?
- 2 What are the civilizational stage, environmental phase, and state of the task process? Do not substitute a version number.
- 3 Which capability is actually needed: sensing, resilience, crisis response, immunity, coordination, interfaces, or long-term subjecthood?
- 4 Who has the authority to decide? Who makes the expert judgment? Who executes? Who provides independent error correction?
- 5 What data, resources, domains, or permissions will be opened? How will exit, return, repair, and termination work?
- 6 Are you citing a specialist core text, historical parameters, a policy formulation, or a research scenario?
- 7 Does this touch a non-regression clause? If so, is there new evidence, or is it merely more convenient to implement?

The minimum contract: version number does not equal authority; capability does not equal permission; output does not equal a decision. High-consequence action must be reversible, allow exit, and be subject to renewed authorization.

Cross-version limits

Cross-version non-regression baseline: Capability does not generate sovereignty; metrics do not generate power. High-consequence action must allow pausing, withdrawal, exit, repair, termination, and renewed authorization. Subjects must not be manipulated. Use only the minimum necessary data. Crisis powers must be withdrawn on expiry. The three goals must remain separately visible and nonexchangeable. Mature systems should be able to step back. Do not import 4.x or Ω into present-day reality.

Current controlled entry: G-03, AGI-COS Controlled Version Map V1.0; Multi-version Continuity Audit V1.1; and G-04, Charter of Constraints and Non-regression.

Civilization Leap Research Group · www.civitas.top · The version map defines only the boundaries of interpretation and use. It does not authorize deployment, data, resources, or public execution.

English translation identity

Zijun Fu | Civilization Leap Research Group. Complete English translation of the Chinese one-page staged entry, P05-AGICOS-011 V1.0, reflowed for readability. English translation prepared 14 October 2026. This is distinct from the existing E3-08 English thematic adaptation.

The named controlled source documents above remain references to the source framework; their inclusion does not establish complete English editions. No deployment or other authority is conferred by this translation.

Personalized Further Reading Guide

Choose the next question after identifying the relevant machine-side track: human responsibility, minimum shared boundaries, or the broader civilizational context.

Human-COS Startup: Judgment and Responsibility Windows Before Irreversibility

Continue from machine-side tracks to human judgment, visibility, and responsibility.

Common Window Protocol: An Open Future and Three Constraints

Read the minimum boundaries that preserve a future open to choice.

Human Civilization's Choices in the AGI Era

Explore the broader account of civilization's choices in the AGI era.

English links lead to corresponding webpages. Some destinations are public introductions or version-family overviews rather than complete English books. Unlinked source references do not establish the availability of a full English text.

[Supplementary English reading index](#)