

AGI-COS Version Map

Distinguishing Civilization-Level Principles, Governance Frameworks, and Software Implementation

v2.0 Formal Submission Edition | July 2026

Zijun Fu

Produced by Civilization Leap Research Group

The source body is v2.0, July 2026. The separate current controlled-reading note is V1.1.

In one sentence

AGI-COS is a civilization-level protocol stack for safe operation. Its foundation sets civilizational principles that must not be breached; its middle layer specifies governance, audit, and crisis interfaces; and its upper layer may be implemented by different software systems and institutions. It is neither a single software product nor an operating system in which AI replaces human governance.

1. Three-Layer Structure

Layer	Question addressed	Typical content	Distinction to preserve
Principles	What must not be breached in the post-AGI era?	Ultimate human decision authority; reversibility; anti-closure; declarations of unknowns; the Three-Goal Gate	Not a product feature list
Governance	How can multiple actors jointly manage risk and responsibility?	Permissions; audits; appeals; circuit breakers; rollback; responsibility anchors	Not rule by a single institution
Implementation	How are specific systems put into practice?	Sensing; logs; model interfaces; policy simulation; software and hardware control	No implementation may rewrite the principles

2. Logic of Version Development

Stage	Core task	Status
L0 Conceptual layer	Define the civilization-level operating system and the role of AGI as an amplifier	Concepts and principles
L1 Constitutional layer	Establish ultimate human judgment, reversibility, and anti-closure as firm baseline constraints	Higher-order constraints
L2 Perception audit layer	Confidence, provenance, declarations of unknowns, logs, and responsibility anchors	Governance/technical interface
L3 Civilizational immunity layer	Identify structural distortion, cascading risk, and the flattening of meaning	Cross-system governance

Stage	Core task	Status
L4 Common Window layer	A minimum crisis-communication interface across institutions and actors	Compatibility and reduced misjudgment
L5 Experimental implementation layer	Test software, hardware, and organizational implementations in bounded settings	Must comply with phase discipline
L6 Long-horizon outlook layer	Safe operation across regions, civilizations, and forms	Open research, not an obligation at the launch stage

3. Phase Boundaries

- At the launch stage, establish only minimum interfaces such as final human review, logs, declarations of unknowns, pause, and rollback.
- At the initial stage, sensing and audit modules may operate in limited settings, without giving the system authority to determine civilizational goals.
- Only at the transition stage should cross-institutional interoperability, shared responsibility, and more complex Common Windows be considered.
- An all-domain civilizational operating framework for the stable stage is a long-term direction. It must not be presented as a unified system that can be deployed today.

4. Questions for Review

1. Which elements are civilizational principles, which are governance design, and which are technical implementation?
2. How can different institutional environments accept shared principles without requiring uniform values?
3. How can software implementers be prevented from changing higher-order rules through technical details?
4. How can cross-system circuit breakers be designed without causing new cascading disasters?
5. What evidence is sufficient for an implementation to move from an experimental version to a higher phase?

Current Controlled Reading Note

P05-AGICOS-002 | V1.1 | Version Selection and Governance Control

The original version text is retained. This page explains only its current applicability, boundaries, and further reading; it does not silently rewrite the historical text.

What this track addresses

When AGI-COS has multiple versions, stages, presentation drafts, and technical drafts, how can readers determine which document applies to which task, rather than treating version numbers as a sequence of upgrades?

What remains applicable

This is the current controlled governance entry point. It identifies version families, historical parameters, conflicting mechanisms, and boundaries for use; it does not replace the text of the specialist versions.

What is not automatically inherited

There is no single highest version. Versions may be used in combination, but their authorities, historical parameters, and automatic triggering conditions must not accumulate through version stacking.

Continue reading

[One-Page Guide by Stage](#) | Enter the version family through a real-world problem and its stage.

[AGI-COS 1.0](#) | Enter the crisis operations track.

[AGI-COS 1.1](#) | Enter the perception audit track.

General principle: AGI-COS has no highest version that covers every task. Versions may be used together; their authorities must not accumulate.

[Version family](#)

[One-Page Guide by Stage](#)