

AGI-COS Global Release Package v1.0

Global Minimum Core Release Package | Complete Text | 2026

Zijun Fu

Produced by Civilization Leap Research Group

AGI-COS GLOBAL RELEASE PACKAGE v1.0

(Global Minimum Core Release Package · Complete Text)

File 1

COS-CORE

AGI Crisis Operating System

Core Protocol v1.0

Cover

AGI Crisis Operating System (AGI-COS)

Global Core Protocol v1.0

Author: Zijun Fu

Series: Civilization Leap Working Protocol

License: CC BY 4.0

Status: Open Global Protocol

Year: 2026

1. Canonical Definition

AGI Crisis Operating System (AGI-COS) is a human-centered crisis governance operating protocol that uses AGI as a controllable cognitive tool to manage all categories of crises across their full lifecycle.

AGI is not the decision-maker.

Humanity remains the governing subject.

The purpose of AGI-COS is:

Ensure no crisis drives human civilization into irreversible failure.

2. Why AGI-COS Exists

Human civilization has entered an era where:

- Crisis speed exceeds human cognition.
- Risks are systemically coupled.
- Irreversibility becomes common.

Traditional governance models cannot operate at required complexity.

AGI-COS provides:

- Unified crisis language
- Human-AI collaboration rules
- Reversible intervention framework
- Global coordination structure

3. Core Principles

AGI-COS operates under three non-negotiable principles:

1. Human Sovereignty

AGI provides analysis only.

Final decisions belong to human authorities.

2. Reversibility First

When irreversible risk exists, reversible actions take priority over efficiency.

3. Anti-Closure

No system may eliminate humanity's future choice space.

4. System Architecture

AGI-COS functions as a civilization-level operating system:

Governance Kernel

→Authority & Authorization

Perception Layer (AGI)

→Prediction, modeling, early warning

Decision Layer (Human)

→Policy choice & responsibility

Intervention Layer

→Crisis response execution

Recovery Layer

→Stabilization & compensation

Evolution Layer

→Iterative learning through pilots

5. Unified Crisis Indicators

AGI-COS introduces global crisis metrics:

CI — Crisis Coupling Index

Measures multi-system interaction risk.

WUI — Window Urgency Index

Measures intervention timing pressure.

IRI — Irreversibility Risk Index

Measures potential irreversible damage.

CSI — Civilizational Stability Index

Measures societal resilience and trust.

These metrics create a shared international crisis language.

6. Standard Intervention Pipeline

Detect

→Diagnose

→Window Assessment

→Impact Mapping

→Path Design

→Execute

→Recover

→Learn

AGI assists at every stage.

Humans retain decision authority.

7. Intended Users

- Governments
- Cities
- International organizations
- Crisis management agencies
- Research institutions
- NGOs
- Corporate risk teams

8. Expected Outcome

AGI-COS does not eliminate crises.

It ensures crises remain manageable, reversible, and survivable.

File 2

COS-START

Crisis Start Guide

Purpose

Enable any organization to activate AGI-COS within 48 hours.

STEP 1 — Activation Criteria

Activate AGI-COS if ANY condition applies:

- Impact ≥ 1 million people
- Rapid risk escalation
- Decision uncertainty rising
- Crisis spreading across systems
- Public trust declining
- Traditional governance response insufficient

STEP 2 — Establish Three Roles

Human Authority

Responsible for decisions.

AGI Analysis Team

Provides modeling and prediction.

Independent Audit Team

Ensures transparency and legality.

STEP 3 — Produce Four Indicators

Within first operational cycle:

- CI
- WUI
- IRI
- CSI

STEP 4 — Enter Intervention Pipeline

1. Detect anomaly
2. Diagnose system dynamics
3. Evaluate intervention window
4. Map consequences
5. Select reversible strategy
6. Execute with monitoring
7. Begin recovery planning

STEP 5 — Continuous Oversight

- Maintain human decision logs
- Monitor AGI performance
- Trigger shutdown if instability occurs

AGI never executes policy independently.

File 3

COS-PRINCIPLES

Non-Negotiable Principles

Principle 1 — Human Sovereignty

AGI must never replace human governance authority.

Requirements:

- Human sign-off mandatory
- Decision accountability remains human
- AGI holds zero executive power

Principle 2 — Reversibility

All crisis interventions must be:

- Stoppable
- Reversible
- Auditable
- Compensable

Irreversible paths require maximum human control.

Principle 3 — Anti-Closure

No crisis response may permanently eliminate societal alternatives.

Systems must preserve:

- plurality
- auditability
- replaceability
- future optionality

These principles form the constitutional layer of AGI-COS.

File 4

COS-PILOT

90-Day Pilot Blueprint

Objective

Validate AGI-COS safely through limited-scale deployment.

Phase 1 (Day 1-30)

Diagnosis & Tool Validation

- Define crisis scope
- Match AGI capability
- Conduct risk verification
- Establish shutdown protocol

Phase 2 (Day 30-60)

Controlled Intervention

- Apply AGI analytical support
- Monitor indicators continuously
- Adjust strategy dynamically
- Maintain reversibility

Phase 3 (Day 60-90)

Evaluation & Iteration

- Publish pilot report
- Evaluate CI / WUI / IRI / CSI
- Improve tools
- Decide expansion or rollback

Pilots must remain small, transparent, and reversible.

File 5

COS-OPEN

Open Global Declaration

AGI-COS is released under Creative Commons CC BY 4.0.

Allowed:

- Use
- Translation
- Adaptation
- Forking
- Localization
- Institutional adoption

AGI-COS is:

- Non-political
- Non-national
- Open infrastructure
- Humanity-oriented protocol

No entity owns AGI-COS.

Human civilization collectively maintains it.

Current Controlled Reading Note

P05-AGICOS-004 | V1.0 | Crisis Operations

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 a high risk of irreversible harm has entered a crisis window, how should detection, diagnosis, window assessment, impact mapping, path design, execution, recovery, and learning be carried out?

What remains applicable

This remains an active specialist track. It retains the full crisis lifecycle, priority for irreversible risks, independent auditing, decisions by accountable real-world actors, and recovery and learning.

What is not automatically inherited

Historical triggering parameters, including the 48-hour period and the number of people affected, no longer have default force. A crisis label must not automatically expand the real-world sovereign authority of AGI or any system.

Continue reading

[AGI-COS 1.1 | Establish perception, provenance, and declarations of unknowns before making crisis judgments.](#)

[Common Window Protocol | Check the baseline constraints on irreversibility, rollback, exit, and non-manipulation.](#)

[Human-COS Public Explanation | Bring direction, responsibility, and real-world actors into the judgment.](#)

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](#)