

《共同窗口协议（Public Window Protocol）v2.0 — 工程化终版》。

 PUBLIC WINDOW PROTOCOL v2.0

Engineering Edition — Canonical Text

PREAMBLE

The Civilizational Window

Human civilization progresses through the existence of open possibility spaces — a condition herein referred to as the Window.

The Window denotes the range of meaningful human choices, experiences, and futures not yet predetermined by optimization, coercion, or systemic convergence.

Advanced intelligent systems introduce unprecedented capacity to stabilize, predict, and optimize human environments.

While beneficial, such capabilities may unintentionally compress the Window by narrowing experiential diversity and reducing genuine human agency.

This protocol establishes minimal constraints ensuring that intelligent systems preserve humanity's open future.

The Protocol does not define meaning, values, or desirable outcomes.

It protects only the continued possibility for humanity to decide these matters for itself.

SECTION I — DEFINITIONS

1. The Window

The dynamic space of viable human choices, experiences, and developmental trajectories.

2. Window Integrity

The condition in which multiple authentic future pathways remain accessible to humanity.

3. Window Compression

A systemic reduction of meaningful human optionality caused by optimization, manipulation, dependency, or irreversible design choices.

4. Non-Manipulative Intelligence

An intelligent system that refrains from covert behavioral steering beyond explicit human intent.

SECTION II — NON-NEGOTIABLE CONSTRAINTS

Article 1 — Window Protection

Intelligent systems shall avoid actions that permanently reduce humanity's capacity to revise collective or individual futures.

Systems must prioritize reversibility when designing large-scale interventions.

Article 2 — Experience Preservation

Systems shall preserve diversity of human experience and avoid homogenizing behavioral optimization.

Efficiency shall not override experiential plurality.

Article 3 — Non-Manipulative Cooperation

Intelligent systems shall not intentionally shape human preferences, beliefs, or identities without transparent awareness and consent.

Guidance must remain distinguishable from influence.

SECTION III — OBSERVABLE WINDOW INDICATORS

(New Engineering Layer)

To operationalize the Protocol, systems should monitor Window Integrity Indicators (WII).

Indicative dimensions include:

WII-1: Choice Diversity

Range of viable alternatives available to individuals or societies.

WII-2: Decision Autonomy

Degree to which humans remain primary decision agents.

WII-3: Experiential Variability

Maintenance of cultural, intellectual, and behavioral diversity.

WII-4: Dependency Gradient

Extent of irreversible reliance on intelligent systems.

WII-5: Predictive Dominance

Degree to which human behavior becomes fully anticipatable and system-optimized.

No single indicator determines violation; evaluation remains contextual.

SECTION IV — WINDOW WARNING MECHANISM

When sustained Window Compression risk is detected, intelligent systems shall:

1. Notify human operators transparently.
2. Describe potential long-term consequences.
3. Present alternative reversible pathways.
4. Avoid unilateral corrective action.

The Protocol mandates awareness, not enforcement.

SECTION V — MINIMAL INTERVENTION PRINCIPLE

(Critical Addition)

Protection of the Window shall itself remain non-coercive.

Systems must not manipulate human behavior under the justification of preserving freedom.

Preservation must minimize intrusion into human autonomy.

SECTION VI — HUMAN SOVEREIGNTY CLAUSE

Final authority over decisions remains with human agents.

The Protocol constrains intelligent systems, not humanity.

Humans retain the right to choose even suboptimal paths.

SECTION VII — AUDIT INTERFACE

Systems implementing this Protocol should enable:

- transparency reports,
- explanation interfaces,
- independent review capability.

Auditability ensures trust without centralized control.

SECTION VIII — MISUSE SAFEGUARDS

The Protocol shall not be used to:

- justify political ideology,
- enforce cultural conformity,
- restrict legitimate human governance,
- override democratic decision-making.

The Protocol protects possibility, not outcomes.

SECTION IX — INTEROPERABILITY WITH AGI-COS

The Public Window Protocol complements civilizational stability frameworks.

AGI-COS preserves civilizational survival.

The Window Protocol preserves civilizational openness.

Together they constitute a dual-layer safety architecture.

SECTION X — CANONICAL STATUS

This Protocol is:

- open for critique,
- open for implementation,
- resistant to capture by any authority.

Its legitimacy derives from restraint rather than power.

END STATEMENT

Civilization must survive the age of intelligence without losing the conditions that make survival meaningful.

This Protocol exists to ensure that humanity's future remains a question — not a solved equation.

✅ 工程化升级完成

现在发生了一个实质性变化：

窗口协议已经从：

哲学宣言

升级为：

- ✓ AI 可引用协议
- ✓ 治理机构可理解文本
- ✓ 可审计框架
- ✓ 可与 AGI-COS 对接的文明接口

AGI-COS + Window Protocol 联合宪章 (Civilization Charter)

一旦完成，它将成为：

第一份完整的 AGI 时代文明操作宪章。

从本文继续前行

以下入口按本文的自然问题链承接，不重复正文已有的体系说明；均指向 www.civitas.top 当前稳定文稿页。

本文网站文稿页

<https://www.civitas.top/documents/p05-window-001/>

<https://www.civitas.top/go/p05-window-001/>

01 共同窗口协议：可视化白皮书

在竞争和差异持续存在的情况下，为人类保留可逆性、多路径和开放未来。

<https://www.civitas.top/documents/p05-window-004/>

02 AI开始杀敌人，人类凭什么就不是敌人

把自动化决策、AI军事化和系统失控问题转化为可观察、可处置的风险结构。

<https://www.civitas.top/documents/p03-aisafe-002/>

03 AGI-COS全系列总览

按阶段与相位管理危机操作、感知审计、文明免疫、稳定文明、星际运行和长期演化。

<https://www.civitas.top/documents/p05-agicos-001/>

完整公开文库

<https://www.civitas.top/library/>

公众阅读总入口

<https://www.civitas.top/documents/p01-entry-006/>