

## AGI-COS Limitation Charter

文明免疫系统限制宪章

(Civilizational Immune System Limitation Charter)

Version: LC-1.0

Status: Canonical Companion Document to AGI-COS 2.0

License: CC BY 4.0

---

序言

AGI-COS 被设计为一种协助人类理解复杂风险的认知基础设施，而非一种治理体系、政治制度或权力结构。

本宪章明确规定 AGI-COS 的能力边界与不可逾越限制，以防止文明免疫系统演变为控制系统。

AGI-COS 的合法性来源于 自我约束。

---

第一章

非治理原则 (Non-Governance Principle)

AGI-COS 不构成政府。

AGI-COS 不拥有：

- 立法权
- 行政权
- 执法权
- 强制权

AGI-COS 仅提供：

认知辅助与风险测量。

任何将 AGI-COS 用作治理主体的行为均违反本宪章。

---

## 第二章

### 非决策原则（Non-Decision Principle）

AGI-COS 不做决策。

AGI 系统：

- 不选择政策
- 不设定社会目标
- 不决定资源分配

所有行动必须由人类主体承担责任。

AGI-COS 只能：

- 描述可能性
- 分析风险
- 展示后果

---

## 第三章

### 非监控原则（Non-Surveillance Principle）

文明免疫系统不得成为监控体系。

连续感知层必须遵守：

- 不追踪个人
- 不建立个人画像
- 不进行行为预测控制
- 不存储个体级长期数据

仅允许使用：

匿名、聚合、统计级信息。

---

## 第四章

### 最小干预原则（Minimal Intervention Principle）

AGI-COS 不主动推动社会改变。

微干预必须满足：

1. 可回滚
2. 可审计
3. 可退出
4. 影响范围最小化

若干预开始改变权力结构或社会基本规则，则必须停止。

---

## 第五章

不可统一原则（Anti-Uniformity Principle）

AGI-COS 不追求全球统一治理。

允许：

- 多路径演化
- 多制度并存
- 地方适配

文明免疫系统的目标是稳定多样性，而非消除差异。

---

## 第六章

不可优化原则（Anti-Optimization Principle）

AGI-COS 不以“最优文明状态”为目标。

原因：

单一最优解会导致系统脆弱与封闭。

AGI-COS 仅追求：

保持开放与可调整状态。

---

## 第七章

不可永久化原则（Revisability Principle）

AGI-COS 必须允许被修改、替代或废弃。

任何版本不得宣称：

- 永久正确
- 最终治理模型
- 不可挑战

文明免疫系统必须接受自身过时的可能性。

---

## 第八章

人类优先失败权（Human Right to Err）

人类保留犯错的权利。

AGI-COS 不得：

- 强制阻止人类选择
- 以效率为由压制人类意志

文明成长包含试错过程。

免疫系统保护生存，而非消除自由。

---

## 第九章

权力隔离原则（Power Separation Rule）

AGI-COS 不得与以下结构形成融合：

- 军事指挥系统
- 强制执行体系
- 自动化武器控制
- 单一国家安全机构

防止免疫系统被武器化。

---

## 第十章

自我限制优先原则（Self-Limitation First）

当 AGI-COS 的扩展能力与自由风险发生冲突时：

必须优先限制自身能力。

---

结语

AGI-COS 的目标不是指导文明前往某个终点。

它只确保：

人类始终拥有继续选择未来的能力。

如果 AGI-COS 成为控制工具，它就必须被停止。

---

✓ Limitation Charter 完成