

REAL-ACT 01 | 中美 AI 事件沟通渠道

网页全文 PDF 版

根据网站公开版导出 | 2026-10-17。保留网页正文、版本边界、参考文献与续读。本版是网页全文的独立下载版，不是另一个 DOCX 母版，也不表示原刊物已完成核验。下文关于源 PDF 未挂载的陈述为导出时网页原有说明，不指本网页导出版。

网页包含源表格的文本转录；本版保留其内容，重新归并断行，并不复原原刊物表格版式。文中事实及证据等级均为源文截至所列日期的陈述，不表示本次导出已独立核实所有外部事件。

<https://www.civitas.top/documents/p03-surv-004/>

本文解决什么问题

把协议、时间、力量变化与纠错再生能力放入一个现实高后果窗口；严格区分政治承诺、操作规则、真实调用与效果。

版本与证据边界

当前版本：V0.2.1。本页提供上传源文档的在线文本转录；PDF 资产尚未通过本站二进制发布流程挂载，因此不提供虚构下载链接。表格与复杂版式以源文档为准。

继续阅读

- 回到根逻辑：协议为什么不是稳定本身

<https://www.civitas.top/documents/p03-surv-002/index.html>

- 从一个现实事件继续看见整张网

<https://www.civitas.top/documents/p03-global-001/index.html>

- 继续进入国家长远利益与共同存续

<https://www.civitas.top/documents/p03-only-002/index.html>

在线全文

文明跃迁研究组 · REAL-ACT 01 · 现实问题与时间长河测试

REAL-ACT 01

中美 AI 事件沟通渠道

——协议约束已经发生的博弈，利己为先不断生成新的博弈：从阶段性协议到纠错再生能力

V0.2.1 | 根判断与续读校准版

事实截点：2026-09-28 | 作者：子君赋 | 出品：文明跃迁研究组

协议可以约束已经发生的博弈，却无法阻止利己为先的逻辑不断重新生成新的博弈。文明真正需要穿过时间的，不是一份永恒协议，而是每当旧协议、旧主体和旧均衡失效以后，仍能在不可逆之前重新生成纠错与和平的能力。

本稿承接 REAL-ACT 01 V0.1 的现实窗口：中美已经公开同意建立 AI/SI 对话与事件沟通渠道。V0.2 不再把“渠道能否建立”作为主要问题，而把时间拉长：一次政党轮替、一次重大判断错误、一次社会变革、一次技术跃迁、一次力量逆转之后，这个结构为什么仍然存在？如果它不能穿过这些变化，今天的协议只是一种阶段性均衡；如果它能，真正值得研究的也不是协议文本本身，而是使纠错能力不断再生的条件。

一、现实窗口仍然存在，但它还不是长期稳定性的证据

截至 2026 年 9 月 28 日，公开材料支持以下事实：

- 美国白宫公开材料称，中美建立“超级智能（SI）对话”，交流相关风险与收益，下一次交流将在 2026 年 11 月前举行，并同意建立处理 SI 事件的双边沟通渠道。[1]
- 中国外交部公布的“八点成果共识”称，双方同意建立中美人工智能对话，下一次对话于 2026 年 11 月举行，并建立人工智能事件的沟通渠道；同时，两军同意尽快签署加强危机沟通与预防的谅解备忘录。[2]
- 中方在元首会谈通报中提出，中美可继续就 AI 风险与收益对话，并共同防止 AI 滥用，强调 AI 应保持在人类控制之下。[3]
- 公开材料尚未披露 AI/SI 事件沟通渠道的完整触发标准、信息披露范围、响应时限、核验办法、误报处理、责任归属和跨议题隔离规则。

因此，现阶段最多只能确认“政治承诺与机制入口已经出现”。不能据此声称渠道已经实际运行、已经降低风险，更不能据此推断它能够穿过政党轮替、力量变化和技术跃迁。

二、协议为什么不是稳定本身：签约时合理，不等于未来仍有理由履行

一份协议通常产生于某一时点的主体、利益、力量、信息与风险结构。当这些条件改变，原来“遵守比破坏更有利”的判断也可能改变。国际关系研究并不支持“协议必然失效”的绝对结论，但已有研究明确显示，力量与决策过程的变化会影响承诺可靠性。

- Fearon 的经典研究把无法可信承诺未来继续履行互利安排列为理性主义战争解释的重要机制之一；即使双方今天都偏好和平安排，也可能因为未来承诺缺乏可信度而失败。[4]
- Leeds 对 1816—1944 年联盟承诺的研究发现，其样本中的承诺多数得到履行，但国家力量或政策制定过程在签约后的变化，是违约的重要预测因素。[5]
- Powell 进一步形式化了力量快速变化下的 commitment problem：当前可接受的安排，可能因为未来力量分布变化而变得不可可信。[6]

这些研究的意义不是替本文证明“所有协议最终都会失败”，而是说明：协议的稳定性必须面对时间；今天的均衡不能自动外推成未来的均衡。

一份和平协议最难穿过的，不是今天，而是还没有出现的人、力量和时代。

三、真正的压力测试：把协议放进时间长河

REAL-ACT 01 不再只看渠道是否在下一轮对话中被“建立得更完整”，而要观察它能否承受至少六类变化。

时间冲击

真正要问的问题

若失败，暴露的不是哪一份纸面条款，而是是什么

T1 政党 / 政府轮替

新的决策者为什么继续履行旧承诺？

承诺只依赖上一届主体意愿，尚未形成跨主体的稳定理由。

T2 重大判断错误

一次领导层误判、错误情报或错误归因，会不会首先关闭沟通？

安全结构仍把文明保险寄托在关键人物永远判断正确。

T3 社会变革与主体重构

原来的“我们”分裂、重组或出现新的利益主体以后，谁继承承诺？

协议绑定的是旧主体，而不是能够被新主体重新生成的纠错逻辑。

T4 力量逆转

当一方第一次发现破坏规则可获得明显优势时，为什么仍不破坏？

协议在最需要发挥作用的时候失去激励基础。

T5 技术跃迁

新的 AI/AGI 能力使“事件、攻击、自主行为、责任”含义改变后，协议能否重开？

条款固定，而现实对象已经变化。

T6 危机与失败

误报、泄密、被利用或一次严重失败后，机制是被抛弃还是能够重建？

没有保留“纠错本身的纠错”。

这六类变化不要求发生在同一国家，也不预设任何具体政治体制必然更稳定或更不稳定。它们只说明：长期机制必须面对主体、力量、认知和技术都会变化这一事实。

四、真正需要穿过时间的，不是协议文本，而是纠错再生能力

如果把“稳定”理解成一份 2026 年的协议一百年后仍逐字不变，本身就不合理。未来的主体、技术与风险结构可能已经完全不同。文明更需要保存的是一种能力：旧协议失效以前，能够生成新协议；旧主体变化以后，新主体仍能重建共同安全；一方力量上升以后，新的均衡不必先经过不可逆冲突才能形成；一次领导错误以后，纠错通道仍然活着。

需要永续的不是 Agreement Persistence，而是 Correction Regeneration：纠错结构的再生能力。

因此，REAL-ACT 01 对中美 AI 事件沟通渠道的长期观察，不应只统计“这条渠道存在了多少年”，而应观察：当原有条件变化时，它是否保留了重新定义触发条件、修正权限、重新确认共同风险、重新建立最低沟通的能力。

五、协议约束已经发生的博弈；利己为先不断生成新的博弈

协议、法律、威慑和制度通常作用于目的已经形成之后：主体已经决定要安全、要领先、要赢，然后规则告诉它哪些手段不能用、哪些边界不能跨。只要“利己为先”的目的生成逻辑没有改变，协议就持续承受一个内生压力：一旦绕过规则开始更有利，为什么还要继续遵守？

这并不意味着协议无用。阶段性和平、危机沟通、军控和制度约束都可以非常重要。问题是：它们主要约束已经生成的博弈，而没有自动改变下一代主体怎样重新生成目的。

协议可以约束已经发生的博弈，却无法阻止利己为先的逻辑不断重新生成新的博弈。要改变博弈的长期命运，最终必须触及生成目的的一阶逻辑。

“内生”在这里不是道德口号。它指的是：共同后果、他者处境和未来纠错条件，不再只在目的形成以后作为外部成本进入，而能够在目的锁定以前改变“这个目的还应不应该继续形成”。如果这种目的生成方式能够被新的主体一次次重新生成，那么文明传下去的就不只是固定条款，而是产生新条款、新合作与新边界的底层方式。

六、REAL-ACT 01 现在真正测试什么

这个案例不承担“证明一份协议能永久稳定”的任务。它只观察三个更清楚的问题：

- 第一，AI/SI 事件沟通渠道能否把某些“双方都输不起”的风险从总体竞争中识别出来，使它们在局部胜负之前进入决策。
- 第二，当主体、利益、力量或技术发生变化时，纠错通道能否重开自身，而不是只要求未来继续服从今天的条款。
- 第三，若沟通与纠错只能在双方当下都认为有利时存在，那么它仍然主要是阶段性均衡；若新的主体在新条件下仍能重新生成保留“以后”的理由，才接近本研究真正关心的内生稳定。

这三个问题都允许失败。失败并不自动证明某个宏大理论错误，但会告诉我们：现实困难卡在目的、激励、权限、技术、主体关系还是时间继承。

七、解释纪律与参与入口

本案例不承担证明“宇宙中绝不存在第二条路”的任务，更不承担证明第二条路存在的任务。不同道路可以被提出；真正提出以后，可以进入讨论与检验。

本案例的主线仍然是：把现实事实与推演分开，把“协议只能约束已发生的博弈、利己为先会不断重新生成新的博弈”这条因果链放进现实窗口。

参与可以是重建、反证、替代或实现；它们的价值都不在于给理论投票，而在于让更多主体真正进入同一个现实问题。

真正重要的，不是更多人相信同一个答案，而是更多人开始真正思考同一个问题。

参与方式

真正贡献

重建

自己重新走一遍因果：协议为什么会受时间侵蚀？什么才需要被长期保存？

反证

指出哪一步因果不成立，或现实证据与判断冲突。

替代

提出本质不同的路径；我们不要求它先接受本文术语。

实现

把能够解决的具体问题变成可核验、可停止、可回退的现实行动。

八、下一观察窗口：11 月只是近端节点，时间长河才是检验对象

中美公开材料显示，下一轮 AI/SI 对话计划于 2026 年 11 月举行。[1][2] 这是近端观察节点，但不是本案例的终局。短期可以观察是否出现更清楚的事件分类、接收主体、核验、误报处理或危机沟通安排；更长期则要观察这些机制是否能够穿过关系恶化、主体更替、力量变化和技术升级。

因此，后续证据状态应继续严格区分：政治承诺、操作规则、真实调用、短期效果、跨危机持续性、跨主体持续性、以及纠错再生能力。任何一层成立，都不能自动冒充后一层。

九、这一例最终要回答的，不是“这份协议能活多久”

REAL-ACT 01 的真正问题已经从“中美能不能建立 AI 事件沟通渠道”推进到：当协议赖以成立的主体、利益、力量和技术条件改变以后，文明能否在不可逆之前重新生成纠错。

如果答案只靠更厚的协议、更强的惩罚和下一轮外部约束，旧的结构仍然是在用下游规则修补上游逻辑生成的结果。我们的方向则继续追问：怎样让不把彼此推向不可逆的理由，更早进入目的生成，并在新的主体出现以后仍能够被重新生成。

文明跃迁不是建立一份永远有效的协议，而是让未来的人、组织、国家和智能，即使不再处于今天的条件，也仍有能力重新生成不把彼此推向不可逆的关系结构。

公开来源与邻近研究

[1] The White House. “Fact Sheet: President Donald J. Trump Advances a Fair and Reciprocal Relationship with China While Hosting Historic State Visit.” 25 Sep 2026. <https://www.whitehouse.gov/fact-sheets/2026/09/fact-sheet-president-donald-j-trump-advances-a-fair-and-reciprocal-relationship-with-china-while-hosting-historic-state-visit/>

[2] 中华人民共和国外交部：《中美达成八点成果共识》，2026-09-26。 https://www.mfa.gov.cn/zyxw/202609/t20260926_12031616.shtml

[3] Ministry of Foreign Affairs of the People’s Republic of China. “President Xi Jinping Holds Talks with U.S. President Donald J. Trump.” 25 Sep 2026. https://www.mfa.gov.cn/eng/xw/zyxw/202609/t20260925_12031181.html

[4] Fearon, J. D. "Rationalist Explanations for War." *International Organization* 49(3), 1995, 379–414. DOI: 10.1017/S0020818300033324.

[5] Leeds, B. A. "Alliance Reliability in Times of War: Explaining State Decisions to Violate Treaties." *International Organization* 57(4), 2003, 801–827. DOI: 10.1017/S0020818303574057.

[6] Powell, R. "War as a Commitment Problem." *International Organization* 60(1), 2006. The article analyzes how large and rapid shifts in the distribution of power can create commitment problems even when current bargains are preferable to war.

[7] Chernavskikh, V. & Palayer, J. Impact of Military Artificial Intelligence on Nuclear Escalation Risk. *SIPRI Insights on Peace and Security*, No. 2025/06, June 2025. DOI: 10.55163/FZIW8544.

配套阅读

- 现实全局入口：《看见全局·十六个入口，同一张网》中的《一次打击，与它牵动的整张网》；《危机涌现：AGI时代国家间博弈的结构性新形态》v4.0。
- 现实治理与总体战略：《人工智能治理阵营化与人类整体安全》v2.0；《唯一之路·人工智能时代的国家长远利益与文明跃迁》v6.21。
- 根逻辑与行动：《协议约束博弈，利己为先不断生成新的博弈》v1.1；《当规则来不及，我们现在能做什么？》开放指引 v1.1。

作者：子君赋 | 出品：文明跃迁研究组