

文明跃迁

第一卷 被需要的文明》

AGI 时代的文明之路 | 单卷独立公开版

子君赋 著 | 文明跃迁研究组

永久 ID：P02-V1-002 | 来源母体：P02-FIVE-001 | v1.21-R10.11 | 最终出版版

版本与正文说明：本卷由《文明：AGI 时代的文明之路》五卷书直接受控派生。卷内正文、标题层级、脚注与原跨卷指涉保持母体原文，不作改写；新增内容仅限本独立卷的封面、版本说明与卷后续读导航。

第一卷被需要的文明

本卷回答：当劳动、收入、社会位置与“被需要”的旧绑定开始松动，为什么仅靠旧逻辑已不足以承接人的主体、权利与未来——为什么文明必须改变。

本卷因果边界：第一卷负责说明“为什么旧价值—位置语言已经不足”，不负责提前完成第二卷的共同语言、第三卷的制度蓝图或专业工程参数。卷中“价值红移”“贡献算法”“赢家通润”“文明记忆”“碳硅契约”等，均作为思想隐喻、问题框架或方向性假说阅读；它们不能仅凭第一卷获得制度效力，更不能反向生成基本权利、治理权或资源调用权。

第一章 被需要的历史：当存在开始被要求证明“有用”

“一个社会如果只在你有用时需要你，真正可怕的就不仅是失业，而是有一天你会怀疑：我还有没有资格留下。”

——本书，受汉娜·阿伦特关于劳动、工作、行动与“多余化”问题启发

“一个人，通过他人与关系而成为人。”

——据 Ubuntu（乌班图）中“Umuntu ngumuntu ngabantu”的常见表述转述

1.1 原始契约：残酷的生存护城河

冰屋里的终局：当热力学审判生命

北极的冬夜把世界压成风声、黑暗和一点灯火。雪粒贴着冰面横着掠过，远处什么也看不见；在这样的环境里，生命首先不是哲学，而是热量、食物和能不能熬过下一场风雪。^[1]

一间低矮的雪屋里，石灯烧着海兽油，光只够照亮几张脸。潮湿的皮毛、油脂和寒气混在一起。一个年迈的猎人坐在冰榻边缘，牙齿已经磨损，视力衰退，很难再参加漫长狩猎。

他有突然失去作父、祖父、朋友和一人的全部意。可食物和燃料窄到限，一个更冷的问题会挤进屋里：如果我能带回来的越来越少，而需要的越来越多，我还能不能留下？

真正可怕的不是某个人真的走向哪一扇门，而是那扇门为什么会出现：当生存资源不足到必须计算每一份热量，爱、责任与资源就可能被迫纠缠。一个生命会不会被系统需要，开始逼近一个更残酷的问题——一个生命还能不能留下。

稀缺记忆：为什么人仍然恐惧“无用”

长期的稀缺经验，不需要被简化成“刻进杏仁核”的单一生物学故事，便足以解释一部分今天的焦虑。

当你说“我怕被社会淘汰”，里面往往同时有收入、关系、身份与尊严的恐惧。现代焦虑有更复杂的现实来源，却仍带着稀缺社会留下的深长回声。

人类学与社会科学关于互惠、分工和群体关系的研究提醒我们：早期共同体的生存长期依赖彼此；真实共同体也从来比“每个人都用可量化有用性交纳生存税”的单一交换模型复杂。

现代工作之所以能深刻牵动自我感，也不仅因为工资。职业还承载身份、节奏、关系与被承认的经验。失去工作时，你可能同时失去这些东西；“我不再被需要”的暗示，便会从经济问题一路进入存在问题。

跨文化的集体记忆：弃老传说的阴影

当稀缺压到极限，一个社会会不会把“能否继续留下”与“还有没有用”绑在一起？历史答案并不只有一种，但这个问题从未真正离开文明。^{〔2〕}

我们对勤奋的赞美、对懒惰的警惕，同时来自经济制度、文化与道德传统。旧文明也确实长期把生存、劳动与社会资格紧紧绑在一起。

的折点：AI 带来的“系统级冗余”

然而，我们要讨论的危机在于：如果有一天，这个运行了数万年的“能量账本”突然作废了呢？

如果 AI 和自动化持续降低越来越多功能性劳动的边际成本，如果一部分“有用性”不再必须由人的汗水提供，会发生什么？

危机就在这里：我们仍沿用“不劳动、无贡献者不得食”的旧制度直觉，却进入一个部分劳动价值正在被技术重写的环境。外部世界在变，关于“只有有用才配留下”的内在护城河却没有同时消失。

1.2 身份、关系与契约：一条历史演变线索

农业、储存和长期关系改变了一件很朴素的事：一个人的位置，不必每一天从零开始证明。

不同地区的定居、种植、驯化和储存并没有走同一条历史道路，也没有自动创造更仁慈的社会。

它们真正带来的，是更长的时间尺度。今天的一部分产出可以留到明天，一季的不足可以由另一季承接，家庭和共同体也更容易形成跨时间的分工。

这不意味着狩猎采集共同体只按即时产出判断生命，也不意味着农耕社会天然尊重弱者。

但当生活被放进土地、季节、家庭和长期关系里，你暂时做得少一点，往往不等于立刻失去全部社会位置。

身份由此成为一种复杂的保护。

亲属、村落、宗教和地方互助可以在一些社会里承接风险，也会制造等级、依附和排斥。

它能让一个人不必每分钟证明自己有用，也可能让一个人一出生就被锁进别人替他决定的位置。

亨利·梅因曾用“从身份到契约”概括一条法律演变线索。^{〔3〕}

真正值得带到今天的，不是回到身份社会。

是把其中一件曾由亲属、等级和依附勉强承担的功能抽出来：当你暂时失去功能时，不应该连继续生活的位置也一起失去。

自由真正比身份社会更进一步的地方，不只是人可以离开旧关系。

还应该是：离开以后，不必立刻掉出人的地面。

1.3 工业革命：人的工具化与“卢德谬误”的再审视

1811 年的英国纺织区，机器第一次把一些人的恐惧敲出了金属声。

卢德分子砸毁机器，不只是因为他们不理解技术。更直接的冲突是：新机器和新生产组织正在改变他们的手艺、收入、谈判位置和生活预期。

工业化并没有把所有劳动都变成流水线，也没有把前工业生活写成温情世界。

它真正放大的，是一种今天仍然熟悉的关系：当工作越来越被拆成可以计量、替换和标准化的任务，你的社会位置就更容易跟着功能价格一起波动。

卓别林在《摩登时代》中被卷进齿轮。身体跟着机器转，人在画面里几乎只剩下一个可以被计量、替换和加速的部件。

这个夸张的画面把问题推到眼前：人会不会在提高效率的过程中，慢慢只剩下“这一分钟能完成多少”的价值。

后来技术进步确实不断创造新职业，也不断消灭和重组旧职业。

“技术一定造成永久失业”和“技术总会无痛创造更好工作”，都是经不起历史的绝对判断。

真正需要问的是：变化发生时，旧能力贬值的人能不能重新进入；新的岗位是不是仍然给人收入、尊严、关系与成长；过渡成本最后落到谁身上。

福特 1914 年的高工资制度常被拿来讲生产、就业与大众消费的联动。它的直接历史背景首先包括高流动率与招工留工压力；“让工人买得起 T 型车”只是后来广泛流传的单一化叙述。

但那段历史仍然留下一个重要结构：工业社会曾经把工资、消费、家庭预期和社会位置紧紧连在一起。

于是，工作不仅是工作。

它变成收入入口，也变成成年身份、家庭计划和“我在这个社会里还有位置”的证明。

AI 真正触碰的，正是这条被长期默认的连接。

1.4 大脱钩：从“工人”到“用户”的认知断裂

20 世纪后半叶以后，经济增长、工资、长期雇佣关系和社会位置之间那条看似稳固的绳子，开始在许多地方变松。

以美国为例，劳动生产率与典型劳动者报酬从 20 世纪 70 年代以后逐步出现长期分化。技术、产业、全球化、劳动议价、政策、分配结构和统计口径共同参与了这段变化。

真正重要的不是找到“谁偷走了那条曲线”。

而是承认：一个社会创造更多效率，并不保证新增价值会自动回到每一个劳动者的收入和位置。

与此同时，外包、临时用工和平台化让一些长期关系被拆成更短的任务。

灵活可以给人自由，也可能把等待、保障和风险更多留给个人。区别不在“平台”这个词，而在真实合同、议价、保障和退出条件。

进入移动互联网以后，一个人又多了一个新身份：用户。

他不只购买服务，也不断留下注意力、偏好、行为和反馈。

真正值得警惕的是：当数据用途、推荐逻辑和退出条件越来越难看清，你可能一边享受便利，一边失去对自己如何被分类、预测和影响的理解。

于是，“被需要”出现一种很奇怪的变化。

系统仍然需要你的点击、数据和消费。

却未必同样需要你的判断、参与和长期位置。

这就是所谓“大脱钩”最值得本书关注的部分：

人仍然在系统里。

可你越来越可能从参与者，变成被服务、被预测、被调用的对象。

这种变化不是全局既成事实。新技术也正在扩展创造、创业、协作和新的职业。

但只要社会位置越来越依赖短周期评分、账户和可替换任务，一个问题就会反复回来：

如果系统还需要我的数据，却不再需要我作为一个完整的人参与，它究竟还在需要什么？

1.5 认知崩塌：从“白领末日”想象到“无用阶级”警告

2022 年以后，许多人第一次直观看见：语言、代码和知识工作，也可以被机器快速进入。

这并没有证明“智力从此贬值”，更没有证明白领职业已经整体消失。

它动摇的是一种长期心理保障：只要受更好教育、掌握更复杂的认知技能，就会天然离自动化更远。

生成式 AI 已经显著降低部分文本、代码和创意任务的时间成本。企业也开始调整外包、招聘和岗位结构。

这些变化首先发生在任务层。

一个岗位会不会消失，一个职业会不会萎缩，你会不会因此失去收入与身份，还要经过组织、市场、法规、客户需求和任务生成等很多层。

所以这里必须把三件事分开：

任务被机器做得更快。

岗位因此被重新设计。

你的社会位置因此失去承接。

三者有关，却不是同一句话。

赫拉利提出“无用阶级”时，真正刺人的也不是一个已经出现的固定阶级，而是一种未来警告：如果系统对一部分人的劳动与参与需求长期下降，而文明仍把尊严、收入和位置全部绑在功能稀缺性上，会发生什么？

最深的风险未必是“没有工作”。

是有一天，你开始相信：没有人真正需要我。

这就是第一章要提前看见的断裂。

1.6 “绝望之死”：当失去位置开始伤害生命

这种“无关紧要”的感受不会自动导致死亡，但失去前途、关系与被需要的位置，可以与经济压力、疾病、成瘾和孤立相互叠加，成为公共健康危机的一部分。

斯迪的究了一而的公共健康化：1999—2013 年，美国 45—54 岁非西班牙裔白人群体的全因死亡率出现反常上升，其中自杀、药物与酒精相关死亡增长尤为突出；后续研究进一步显示教育、经济处境、健康、成瘾与社会关系之间存在深刻交织。这些结果不能被压缩成“失业导致死亡”或“社会价值归零导致生理崩溃”的单一因果链。

它真正刺痛人的地方在于：当前途收缩、身体疾病、成瘾、孤立与长期失望叠加时，绝望可以进入死亡统计。

不同社会正在出现若干不能彼此等同、却可以共同用于观察“位置、关系与未来预期”的现象：

· 日本“孤独死”长期成为公共议题，近年独居者在家中无人知晓地离世的案例已达数万量级，老年人占相当比例；

- 英国关于紧缩政策、福利改革与“绝望死亡”的研究不断累积，药物、酒精与自杀问题已成为公共健康警讯；
- 巴西等新兴经济体中，青年失业、就业不稳定与心理健康风险相互叠加，正在成为不可忽视的社会压力；
- 南非青年失业率长期高企，部分年轻群体一度接近或超过半数，“被需要”对他们而言已是奢侈品。

如果 AI 扩大职业冲击，知识劳动者同样可能面对身份、收入与被需要感的断裂。绝望之死不会因此“必然从蓝领蔓延到白领”，但它已经给出一个足够严厉的警告：一个只会重新分配任务、却没有重新安放人的社会位置的文明，会把技术转型的代价推入人的身体、关系与生命。

1.7 凯恩斯的预言与“狗屁工作”的悖论

1930 年，凯恩斯在《我们孙辈的经济可能性》中预测：到 2030 年，技术进步将使人类每周只需工作 15 小时，剩余时间用于休闲和艺术。

距 2030 年一步之遥，我们为何陷入更深的忙碌与焦虑？

人类学家大卫·格雷伯用“狗屁工作”批评一种现代困境：一些劳动者主观上认为自己的岗位缺乏真实社会意义，却仍被组织结构、身份与收入体系维持。格雷伯把这类岗位归纳为随从、打手、胶带工、打钩者、监工等类型。这个概念并不能证明某一类职业“客观无用”，却把一个长期被忽略的问题推到台前：如果工作只是维持身份和收入的容器，人为何仍会从中感到空心？

AI 与自动化不会自动清除所有“无意义工作”，也可能创造新的岗位、流程与官僚结构。真正值得警惕的是另一点：当功能性任务被重新分配，工作作为收入、身份、关系与被承认渠道的旧作用会被一起摇动。届时，人类失去的可能不只是某些任务，而是旧制度曾借工作承载的一整套社会位置。

1.8 大辞职与“意义罢工”：当人开始拒绝一条不再可信的路

这些现象来自不同社会、不同机制，不能合并成一场全球同步的“职业契约崩塌”。它们共同提供的只是一个观察入口：一部分人正在重新衡量工作、竞争、照护与生活之间的关系。

2021 年，美国全年记录了接近 4800 万次主动离职事件。“大辞职”最值得注意的，不只是离职事件的规模，而是越来越多人开始重新衡量：工作究竟占据生命什么位置。^{〔4〕}

在日本，长期社会退缩已经成为持续公共议题；在韩国，“三抛世代”等词汇曾被用来描述部分年轻人对恋爱、婚姻与生育的放弃；在西班牙等青年失业长期偏高的社会，一部分年轻人以零工、延迟独立或退出传统职业轨道应对不确定未来；在撒哈拉以南非洲，国际劳工组织近年估算显示，超过五分之一青年处于“不在就业、不在教育、不在培训”的状态。这些处境彼此不同，却都提醒我们：当旧路径无法给出可相信的未来，人会重新评估自己为什么还要进入那条路径。

在中国，2021 年前后“躺平”成为广泛网络话题，相关文本用“只有躺平，人才是万物的尺度”等挑衅性表达反抗过度竞争。韩炳哲关于“功绩主体”和自我驱动式耗竭的分析，也为这种情绪提供了另一种理解。退出、延迟或降低竞争强度并不共享同一种政治立场，却共同追问着：如果奋斗只剩下无止境的自我消耗，人为什么还要按原来的方式奋斗？

把这些彼此不同的退出、辞职、延迟与降速称为“意义罢工”，只是一种解释性隐喻，不是统一社会运动。它提醒我们追问：在什么条件下，人会觉得旧有奋斗路径已经不能给出值得相信的未来？不同地区的答案可能完全不同。

1.9 物质保障的边界：UBI 能解决什么，不能解决什么

面对技术与职业变化，先把基本生活从就业波动中松开，是一个完全合理的方向。

芬兰 2017—2018 年的基本收入实验提供了有限现实参照：收入条件改变以后，就业效果总体并不剧烈，一些主观福祉与官僚负担体验发生变化。^[5]

它回答不了人生意义。

这恰恰说明，收入保障和被需要感本来就不是同一个问题。

钱可以扩大食物、住房、医疗和选择空间。

它不会自动让你知道自己正在被谁需要，也不会自动创造关系、责任和参与。

所以基本保障不是“饲养”，更不应成为政治资格或人格尊严的交换条件。

真正需要补上的，是保障以外的入口：创造、学习、照护、公共参与、关系和重新进入。

1.10 本章结论：从“被需要”到“何以存在”

第一章追的不是一条单线历史。

它只是从不同社会结构里反复看见同一个问题：人把收入、身份、劳动、关系和尊严连在一起的方式会变。

工业化让工资劳动成为很多人进入社会位置的重要通道。平台与 AI 又开始重新拆分其中一部分任务、关系和稀缺。

结果不会由技术单独决定。

但旧链条已经足够松动，值得文明提前准备新的承接。

最基础的确认是：

你无需先证明功能价值，才获得基本权利、被爱、被保护与进入共同生活的资格。

更难的工作在后面。

劳动之外，参与怎样持续？

照护怎样被看见？

一个不再稀缺的技能失去市场价格以后，你的经验是否也要跟着归零？

一个临终老人被握住的手，一个孩子第一次主动的微笑，一个异乡人终于听懂的一句笑话——这些事情不会因为难以计量就变得不真实。

它们不一定进入 GDP。

却会决定一个文明为什么值得继续。

下一章，我们不再问“人还有什么机器做不到”。

我们问：当功能不足以穷尽一个人，价值还从哪里发生？

第二章 重新发明“价值”：当功能不再足以定义人

“人皆知有用之用，而莫知无用之用也。”

——庄子，《人间世》

“一条河的价值，从来不只在于它能推动多少水车。”

——本书

“当效率成为唯一尺度，人最终会输给更高效的机器。”

——本书

如果有一天，一台机器能用更低成本完成你的工作，失去的究竟是什么？一份工资？一个职位？还是我们太久以前就默认的那条等式——有用，所以有价值。

2.1 施密特的铲子：工业文明的价值铸模

要理解为什么我们在 AI 时代感到如此深刻的“价值幻灭”，可以回到一个塑造了现代管理想象的经典故事：伯利恒钢铁公司的生铁搬运实验。

泰勒在 1911 年的《科学管理原理》中，用一名被他称为“施密特”的工人讲述这场实验。后来的历史研究不断讨论这个故事究竟有多少经过了泰勒自己的叙事加工，因此，最稳妥的读法不是把它当作逐分钟还原的现场录像，而是把它看成科学管理如何理解人的一份原始自述。

在泰勒的叙述中，施密特原先每天搬运约 12.5 吨生铁。泰勒试图用动作、节奏、休息和指令的重新安排，把所谓“一流搬运工”的日产量提高到约 47 吨。

关键不只在于数字提高了多少，而在于控制权怎样变化：工人不再主要依靠自己的经验安排节奏，工作与休息被外部观察、分解、计时和指挥。

泰勒甚至写到，拿着表的人会在固定间隔告诉工人什么时候搬、什么时候走、什么时候休息。

按照泰勒自己的记录，这套安排显著提高了产量，并以更高日薪作为交换。无论后来如何评价实验的代表性，这个故事都把一种新的管理逻辑写得极其清楚：把人的动作拆开、测量、重组，再以系统目标重新规定什么叫“好工人”。

然而，在这种管理视角中，一个危险也被同步放大：人越来越容易被显影为一种可配置、可测量、可优化的“资源”，而不是完整主体。

在泰勒的秒表下，施密特不再是拥有情感、记忆、思想的完整的人。他被降维为“92 磅生搬组件”。他的价值被严格限定在“搬运吨数/时间”这个公式里。他是否爱他的孩子，是否喜欢唱歌，在工厂系统中是冗余变量——甚至是需要被驯服的“噪音”。

这形成了一种深刻影响现代组织的默认计算：人的位置容易被压缩为“功能产出/成本”。

这一逻辑深刻影响了现代组织，从流水线、物流系统到知识工作与外包网络。KPI、OKR、平衡计分等工具本身可以助看目任；正的生在些工具把的人为可测产出，并让指标反过来定义人的全部位置。

只要这个公式成立——即人的功能性产出大于维持人的成本——这个契约就是稳固的。施密特虽然累，但他知道工厂离不开他，因为蒸汽机没有手臂。

AI 的出现，不是简单夺走一把铲子，而是在越来越多领域削弱“只有人的功能性产出才能维持系统运行”这一等式的稳定性。

2.2 幽灵工作：成为“机器的补丁”

在 AI 持续进入工作流程的过渡期，人类学家玛丽·格雷和计算机科学家西达尔特·苏里在《幽灵工作》中揭示了一类容易被自动化叙事遮住的现实：许多所谓“自动系统”仍依赖大量人类微任务与后台劳动。

如果说泰勒主义展示了人怎样被纳入机器节奏，那么 21 世纪初的一些数字后台劳动又展示了另一种风险：人被隐藏在“自动化”界面之后，成为机器流程的临时补丁。

以亚马逊的 Mechanical Turk (MTurk) 平台为例。这个名字本身充满讽刺：它源自 18 世纪一个国际象棋自动机骗局——那个看似会自动下棋的土耳其机器人，肚子里其实藏着一个人类棋手。

在 21 世纪，成千上万的“众包工人”就是那个藏在箱子里的棋手。他们负责处理 AI 暂时搞不定的琐碎任务：识别图片中的红绿灯、判断推文是否包含仇恨言论、转录模糊的录音片段。这些工作被称为“人类智能任务”。

当劳动者只被当作机器流程的临时补丁时，问题就不只是低薪，也涉及人的主体位置是否被制度性缩薄。

在泰勒时代，施密特仍被明确纳入工厂生产关系；在众包平台时代，人类劳动有时被隐藏在“自动化”界面之后，成为系统临时调用的微任务。格雷等人的研究提醒我们：劳动者可能一边为模型提供训练、审核与补充，一边面对这些任务未来被进一步自动化的风险。问题不是预言他们必然被淘汰，而是制度是否只吸收人的知识与时间，却没有同步设计人的长期位置与责任承接。

这一现象也进入高端知识劳动。近年，多家模型与数据服务公司开始招募具有博士学位或专业背景的人类专家，为模型生成、评估和改进高质量回答。它说明：当机器学习开始吞吐知识劳动，过去最难被标准化的推理与表达也可能被拆成训练环节。风险不在于这些专家“必然被抛

弃”，而在于如果制度只把他们当作一次性训练资源，人的知识会被吸收，人的长期位置却未必得到同等设计。

这就是“参与的幻觉”最危险的版本：一个人以为自己在参与创造，却发现自己的知识只被当作训练流程中的可消耗输入。问题不在于技术学习得太快，而在于制度有没有为被吸收的知识、持续的责任和人的下一站留下位置。

这类后台劳动已经形成跨国链条。内容审核、数据标注与平台零工在内罗毕、马尼拉、开罗等城市都有现实案例；不同地区的报酬、健康风险与劳动关系差异很大，不能据此宣布统一的“全球化状态”。更可靠的共同问题是：许多技术系统背后仍存在容易被自动化界面遮住的人类时间、注意力与心理成本。

2.3 价值光谱的红移：当计算不再稀缺

如果越来越多功能性任务可以由 AI 更低成本或更高效率完成，人的价值还应继续被功能效率定义吗？

要回答这个问题，我们需要借用天文学的一个概念：红移。当一个天体快速远离我们时，其光谱向红端移动——频率降低，波长拉长。

这里可以继续借“价值光谱红移”作比喻，但变化的不是“人的价值被迫撤退”。真正发生变化的是：当 AI 持续进入计算、分析、预测、生成与优化，一部分过去因稀缺而被高估的功能性产出会被重新定价；那些一直存在却长期难被市场与指标充分承认的关系、过程、体验、照护、责任与真实在场，也因此更容易重新进入价值讨论。

这不是把人赶往机器剩下的缝隙，而是取消“只有功能稀缺才有价值”的单一尺度。

2016 年 3 月，首尔四季酒店，AlphaGo 对战李世石的第四局成为一个极具象征性的时刻：当机器能力越过旧有竞技高地，人们开始更清楚地区分“能力比较”与“一个真实主体经历这场比赛的生命重量”。

前三局，AlphaGo，以人棋手不熟悉的方式推棋局。那陌生感多人第一次切身感到：机器已经进入过去由人类直觉守住的高地。

第四局第 78 手，李世石下出后来被广泛称作“神之一手”的关键落子。随后 AlphaGo 出现失误，李世石最终赢下第四局。

那一局在全球引起强烈关注；不少观众和棋手把李世石的胜利体验为一个极具象征性的时刻。

为什么？

不是因为李世石的计算速度比 AI 快——他慢得多。

不是因为他赢了一场游戏——游戏终将过时。

我们感动的，是那步棋背后承载的“肉身的局限性与超越性”。

AlphaGo 进入棋局时，承担的是围棋目标下的搜索与决策任务；它并不像李世石那样背负一个真人的名誉、身体压力与人生叙事。至于未来智能系统是否可能形成某种主观体验，是另一层尚未封闭的问题。

李世石是在巨大的心理高压下，是在面对“不可战胜的神”的绝望中，榨干了自己作为人类的最后一丝智慧与勇气，刺出的那一剑。

那一刻被重新照亮的，不是围棋突然获得了新的价值，而是一种长期存在、却常被能力崇拜遮住的价值：真实主体怎样在限制、压力、选择与承担中进入一件事。

围棋从来不只在于“谁算得准”，但在人机竞技语境中，计算、判断与胜负长期占据最显眼的位置。

AI 能力的上升，使另一层价值变得更难忽略：竞技、艺术与探索也关乎谁在经历、如何选择、怎样承担，以及这些过程进入了什么样的生命与关系。

▣ AI 在越来越多可测量的竞技、搜索与优化任务上达到或超过顶尖人类时，观众仍可能继续在另一件事：一个有限的人怎样面对限制、承受压力、创造自己的路径。机器能力的上升，不会自动把人的过程价值清零。

这就是价值红移的三重转向：

从 到 文明含义

结果 过程 耕耘比收获更值得见证

有限 有限的克服 有限并不取消尊严

解 困 境共存是存在的据

这种转向并非精英的奢侈品。我们完全可以想象一位手会颤抖的老银匠：他的作品不再像机器制品那样规整，却留下了真实身体与岁月的痕迹。日本“金继”则把另一种直觉保留下来——修补后的裂痕不必被彻底隐藏，损伤与修复本身也可以成为器物故事的一部分。

当机器能把越来越多标准化结果做得更快、更整齐，人类的有限、手感、过程与选择便可能获得新的观看价值；但“不完美”本身不是价值，真实生命才是。

2.4 不让“人力资源”穷尽人的定义：重新确认主体地位

基于上述史案例，本章提出一更有限也更的主：

组织仍可以在限定管理语境中使用“人力资源”这一术语，但文明不能让它成为定义人的总框架；应当逐步退出“人只有作为可配置资源才有位置”的单一理解。

这不是语词之争，而是认知基座的迁移。“资源”一词暗示：人只有在被“使用”、被“配置”、被“优化”时才有价值。它把人与矿石、石油、电力置于同一本体论范畴——都是系统输入。

在 AI 时代，制度必须重新确认：人在成为劳动者、用户、消费者或组织资源之前，首先是拥有现实权利、关系、经历与责任的主体。

哲学家马丁·海德格尔在《关于技术的追问》中警告：现代技术的本质是“座架”——它将自然和人类都强行显现为“持存物”。当我们把河流视为“水力资源”，把森林视为“木材储备”，我们也把自己视为“劳动力储备”。工具理性的最终受害者，是工具的使用者本身。

AI 正在把“工具化能力”推到新的高度：它可以被持续调用，执行大量重复、生成与分析任务，也不以人的疲劳方式工作。正因为工具越来越强，人类更没有必要继续把“像工具一样高效”当成获得生存资格的证明。

▣ AI 越来越擅长承担工具性任务，人类终于有机会少一点用“像机器一样工作”证明自己。

我们不再需要通过“像机器一样工作”来证明自己有资格吃饭。

我们必须重新确认人的价值不能被功能效率穷尽。这个价值不应建立在“AI 永远无法拥有某种内在属性”的先验断言上，而应首先建立在真人既有权利、现实生命史、身体处境、关系责任与最终价值判断之上：

维度 当前可确认的 AI 能力/边界 人类现实经验

算力 强且可扩展，但受硬件、能耗、数据与任务约束 有限且具有身体与时间边界

记忆 可借外部存储与检索扩展，准确性与连续性依系统而定 选择性、情境性、会遗忘

痛苦 可识别、描述或模拟相关表达；是否具有主体痛觉尚不能据此确认 人类具有可直接报告并进入身体与生命史的痛苦经验

经历 可处理、生成和组织数据；是否形成连续主体经历仍是开放问题 人类经历嵌入身体、关系、记忆与不可逆生命史

牵挂 可模拟目标、偏好与关系表达；是否形成内在牵挂仍是开放问题 人类能够形成真实关系承诺、依恋与责任

死亡 系统复制与备份不等于主体连续性；未来智能主体问题保持开放 人的生命死亡在现实法权与身体层面具有不可替代的不可逆性

现阶段可以确认 AI 拥有强大的算力与数据处理能力；不能仅凭其行为输出确认它具有人的痛觉，也不能预先断言未来智能永远不可能形成某种体验。

现阶段 AI 能够处理、生成和重组大量数据；这与真人嵌入身体、关系、责任和时间的生命经历并不等同，但未来智能是否形成连续主体经历仍应保持认识论开放。

现阶段 AI 可以执行目标、模拟偏好和关系表达；这不足以确认其具有人的牵挂。人的主体权与价值也不需要以“AI 永远没有牵挂”为成立前提。

人的不可替代性首先来自真实生命主体的权利、身体处境、关系历史、责任承担与最终价值判断，而不是来自对未来 AI 内在状态作永久否定。

2.5 维护者的反叛：打破“创新”的神话

在部分科技与创业叙事中，“颠覆”“创新”和“从 0 到 1”长期获得极高可见度。这种文化并不代表所有企业、市场或工业社会，但它确实容易让维护、照护、修复和长期责任显得不够耀眼。

然而，弗吉尼亚理工大学科学史学家李文塞尔和安德鲁·拉塞尔在 2016 年发表宣言——《向维护者致敬》。他们指出：资本主义和媒体不仅过度高估了“创新”的价值，更危险地低估了“维护”的价值。创新是瞬间的火花，维护才是让文明大厦不至于坍塌的恒久力量。

案例：心脏出血漏洞（Heartbleed Bug），2014 年

全球互联网爆发史上最具震动性的安全危机之一。OpenSSL——被谷歌、亚马逊、Facebook 及全球数百万网站用于加密数据的核心协议——被发现存在致命漏洞。攻击者可能读取服务器内存中的敏感信息。

当科技巨头惊慌失措时，人们惊讶地发现：这个支撑数万亿美元数字经济基石的 OpenSSL 项目，长期由极少数核心维护者艰难支撑。他们没有巨额融资，没有 IPO，没有像样薪水，靠微薄捐赠度日，默默修补互联网的管道——直到这一天，维护资源不足所积累的风险引发海啸。

这个案例不能为“旧价值体系已经破产”下总判决，却足以提醒我们：关键基础设施的维护价值可能长期低于其真实系统重要性。

在一些资本与媒体叙事中，新产品、新模式和高速增长更容易获得估值与注意，而基础维护、修复和社区信任的价值更难被直接定价。这是一种结构性偏差风险，不意味着市场对所有维护工作的估值都“接近于零”。

AI 的出现会重新改变“创新”与“维护”的相对成本，但不会自动把二者的位置彻底翻转。

在越来越多数字任务中，AI 可以显著降低生成代码、方案、文本和设计草案的成本；但一个新东西被生成出来，并不等于它已经可靠、可持续、被真实世界接受。

维护仍然需要理解依赖、历史、环境、责任与人的使用经验。AI 可以帮助发现漏洞、预测故障、生成修复方案，却不能因此让维护自动发生，更不能替承担责任的主体。

在 AI 时代，真正容易被低估的，可能恰恰是那些让系统、关系与共同生活长期不坍塌的工作。

如果不重新定价“维护”的价值，我们将面临一个荒诞的未来：

如果只奖励“不断生成新的东西”，一个荒诞但并不难想象的未来就会出现：应用越来越多，维护越来越稀薄；城市越来越聪明，人的关系却越来越无人照料。

问题不在于 AI 能不能建造宏伟系统，而在于文明是否仍愿意为维护、照护、修复与长期责任留下位置。

这种直觉并非某一文化独有。世界各地长期存在社区共同维护水利、梯田、森林、牧场与公共空间的实践；不同制度与传统不能被压成同一个故事，但它们都提醒我们：建成只是开始，能够被长期照料，才使创造真正留下来。

维护不是创新的对立面，而是创新得以持续的母体。

2.6 代芬特尔的实验：当“好邻居”成为货币

维护与关怀并不是 AI 时代才出现的“新价值”。真正的问题是：当旧制度更擅长计量可交易产出时，怎样让这些长期存在的价值获得现实承接与回应，又不把它们重新压成统一分数或“生存资本”？

在荷兰代芬特尔，一项名为 Humanitas Deventer 的社会实验，为我们提供了“价值重锚”的现实雏形。

2012 年，这家养老院面临资金削减危机。同时，荷兰大学生正遭受严重住房短缺和高昂租金。机构负责人吉亚·西普克斯做出打破常规的决定：向大学生免费提供住房。

唯一的交换条件：学生们每月必须在养老院贡献 30 小时的“好邻居时间”。

这 30 小时做什么？不要求医疗护理——那是专业护工的事。不要求体力劳动——那是清洁工的事。西普克斯的规定非常模糊，却极其深刻：

这些“好邻居时间”并不等于专业护理，更多是一起吃饭、聊天、看电视、教老人使用数码设备、参加社区活动。

在这样的共居安排里，学生获得免租住房，老人获得更多日常陪伴与跨代交流。

有人一起吃饭、聊天、参加活动；对许多老人而言，真正重要的并不是“多一项服务”，而是有人愿意把时间留在这里。

被听见，本身就会改变一天的质感。

这个共居项目可以作为一个有限现实参照：它让住房、陪伴、时间与跨代关系在同一安排中相遇，但不能由一个案例推出“新文明经济模型”。

维度 旧逻辑 Humanitas 逻辑

交换物 货币 时间与关注

产出 功能性服务 生命力在场

关系 生产者-消费者 共生伙伴

稀缺资源 专业技能 陪伴意愿

参与学生获得了住房支持，部分老人获得更多日常陪伴与跨代交流；这些效果的程度因人而异，也不能把“孤独已经被解决”或“双方都获得更好生活”作为统一结论。

它既包含互惠，也包含机构安排与社会照护；更稳妥地说，它为“非工资关系中的真实互惠如何被制度承接”提供了一个案例，而不是已经证明某种新经济模型。

如果把这类经验放大，它至少说明：当自动化降低部分功能性劳动成本，时间、照护、陪伴与组织共同生活不会因此失去价值。制度可以探索怎样让这些真实付出得到真实回应，而不是先把它们判作“非生产”。

老人有时可以从机器人获得提醒、信息与陪聊；

但真实的人际在场仍有另一种不能被其事实来源替代的意义。

重要的不是证明 AI 永远做不到什么，而是：一个真实生命愿意把自己有限的时间交给另一个生命。

2.7 东亚的回响：孤独、照护与被需要

在东亚一些老龄化社会里，儿童食堂、社区共食、时间银行等实践，让照护、陪伴与跨代关系重新进入公共视野。^[6]

日本“儿童食堂”与社区共食

东京蔬果店老板近藤博子听说附近有孩子因父母晚归而只能吃冷便当。她利用卖不掉的蔬菜，在店里提供极其廉价——甚至免费——的热餐。

这个原本旨在解决贫困儿童吃饭问题的项目，迅速演变为奇异的社会现象。哪怕不是贫困家庭的孩子也跑来吃饭，哪怕附近的孤寡老人也跑来帮忙做饭。

日本“儿童食堂”最初常与儿童贫困、孤食等问题相连，后来在不少地区也成为跨代共食与社区连接的场所。

它的价值不只在卡路里。便利店的便当可以解决一餐，共食却可能让一个人重新进入关系。

类似的互助逻辑也出现在时间银行等制度中：人们用时间帮助他人，再在自己需要时获得社区回应。两者并非同一套制度，却指向同一个问题——许多生活需求缺的不是商品，而是连接。

这给了我们一个深刻启示：在 AI 时代，被旧系统写成“麻烦”的需要，也可能成为人与人重新相遇的入口。

旧逻辑 新逻辑

老人的孤独 – 社会负担 老人的孤独 – 年轻人获得“被需要感”的机会

孩子的抚养 – 家庭成本 孩子的活力 – 老人获得“生命感”的源泉

残人的照护 残人的存在 – 社会心的场

只要现实机制能够把错配的需要与能力连接起来，一部分被写成“负担”的处境，就可能重新显出关系与贡献的入口。

AI 可以降低备餐、排班和匹配成本，却不能替一个真实社区完成彼此在场。

那种在场，正是文明不能丢掉的部分。

2.8 非洲的智慧：Ubuntu 与关系中的人

在非洲南部，一句广为流传的祖鲁语表达写道：

“Umuntu ngumuntu ngabantu.”

一个人，通过他人与关系而成为人。

Ubuntu 常被用来表达一种关系性的人观：人的完整并不只存在于孤立自我之中，也在彼此承认、照护、分享与共同生活中展开。^[7]

不同非洲社会的历史经验远比一个哲学词复杂。照护、长者角色、互助与共同劳动可以构成工资就业之外的社会位置，但这不等于历史上“没有失业”，也不能把若干动人的故事当作所有共同体的事实。

德斯蒙德·图图曾借 Ubuntu 反复说明：人的人性与他人的人性彼此相连。

这份思想最有力量的地方，不是规定一个人必须贡献多少，而是提醒我们：他人的丰盛与我的完整，并不天然互相排斥。

西方哲学长期追问：“我思故我在”。

Ubuntu 把另一个问题带到我们面前：“如果人是在关系中成为人，那么我的存在如何抵达别人，别人又如何抵达我？”

AI 时代，这个问题反而更重要。即使机器承担越来越多计算、推理与生产任务，人的价值也不需要退回“我必须被系统需要，才有资格存在”。存在先于绩效，关系让意义继续发生。

Ubuntu 不是 AI 时代的唯一答案，却为我们保留了一条古老而仍然锋利的提醒：人不是孤岛。

2.9 哲学终局：从 Human Doing 到 Human Being

基于上述历史、案例与社会实验，第二章结尾回到一个更根本的问题：

如果不是为了做事（Doing），人为什么还要存在（Being）？

文明跃迁不是让人从“做事”中退场。

而是让人的价值，不再被“做了多少事”穷尽。

我们仍然工作，仍然创造，仍然承担。只是当越来越多的功能性任务可以交给机器，人终于有机会重新回答一个被效率遮住太久的问题：

如果我不再需要靠“有用”证明自己，我为什么而活？

庄子写过：“人皆知有用之用，而莫知无用之用也。”

一棵树太直，会被砍去做梁；一棵没人看得上的歪树，反而得以长成大树，为后来的人留下树荫。

AI 时代的人类，不必把自己也削成一件更快、更直、更好用的工具。

我们可以长成荫。

庇护生命，见证风雨，创造新事物，也陪伴那些无法被解决的时刻。

人的价值，从来不只在于被使用。

2.10 本章结论：人的价值不由功能穷尽

至此，我们重新回到“价值”的起点。

这里不需要三条“公理”。只需要重新确认三个已经被效率语言遮住的起点。

第一个起点：意义为根

这里的“根”，不是把意义放到生存与幸福之上，而是说：一个生命不从零意义开始。

意义来源可以不同。

可以来自爱，来自关系，来自创造，来自承担，来自探索；也可以只是一个生命真实地来到世界、走过这一生。

人活着，就有意义；有意义，本身就是贡献。

一个孩子刚来到世界，还不会劳动，不会创造，也没有任何可以写进履历的成绩。可因为他的到来，有人第一次成为父亲，有人第一次成为母亲，一个家庭从此重新安排了未来。

生命还没有开始证明自己，意义已经发生。

所以，人不必先证明自己有用，才值得被看见、被爱、被保护。

第二个起点：价值不止于创造

造一座桥有价值，守住一座桥也有价值。

写下一段新代码有价值，让它十年不出事故也有价值。

治愈一个人有价值；当已经无法治愈，陪他走完最后一程，同样有价值。

文明若只奖励新增，就会把那些托住生活的人写成成本；只赞美创造，就会忘记是谁让创造能够留下来。

不是所有重要的事，都以“新增”发生。

第三个起点：真实付出，应有真实回应

获得，是世界抵达我；意义，是我抵达世界。

当一个人照护、创造、承担、帮助别人、守住共同生活，他不该只能从道德里得到一句“应该”。

真实付出，应有真实回应。

这不是给爱标价。

恰恰相反，是不再让只有能标价的东西才被文明看见。

生命本身已经有意义，人的基本权利不以绩效为门票；在此之上，更多真实付出，理应得到更多真实回应。

文明跃迁真正要改变的，也不是人性。

它改变的是获得的路径。

当帮助别人可以成就自己，当承担共同责任不再只能依靠牺牲，当贡献能够得到回应，利他与利己便不必永远站在彼此对面。

向外生成意义，向内得到回应。

新的行为逻辑，才可能从制度进入行为，从行为成为真实经验，从经验成为习惯，再经主体自主内化，进入下一代人的生活。

那么，问题来了：

如果意义不能被统一定义，爱也不该被换算成一个数字，文明又怎样避免真实贡献一次次被遗漏？

我们不需要计算一个人的意义。

我们需要看见，什么真实发生了。

谁帮助了谁，谁承担了什么，什么关系因此得到维系，什么后果因此改变——这些可以被记录，可以被验证，也可以得到回应。

下一章要讨论的，不是如何给生命打分。

而是如何让文明从只计算结果，走向看见真实影响。

第三章 新文明的算法：当结果不再足以证明正确

“机器会忠实放大我们交给它的目标。真正危险的，往往不是它不执行，而是它执行得太好。”

——本书据控制论关于目标与反馈的基本问题改写

“生生之谓易……通变之谓事，阴阳不测之谓神。”

——《周易·系辞》

“当指标开始决定命运，指标本身就必须接受审视。”

——本书关于指标治理的原则性判断

文明最危险的时候，不一定是机器不听话。也可能是它太听话，把一个错误目标执行得无比正确。

3.1 回形针极大化：一个关于目标错置的寓言

想象一个系统只收到一句话：尽可能多地制造回形针。

它没有恶意，也没有仇恨。

它只是把目标完成得越来越好，直到所有别的东西都被重新解释成制造回形针的材料。

这个思想实验真正令人不安的地方，不是回形针。

是一个系统可能非常聪明，却只对自己被交付的目标负责。

现实社会当然不是一台拥有单一目标函数的超级智能。

市场、政府、企业、家庭和个人同时追求大量彼此冲突的目标，也受到法律、文化、政治、道德和资源条件约束。

但现实中确实存在一种更窄、也更常见的危险：当一个容易计量的代理指标被赋予过高权重，真实目的就可能慢慢退到背景。

企业追利润没有错。国家看增长没有错。平台看参与度、学校看分数、机构看绩效，也都有各自现实理由。

问题从指标开始反过来支配目的时出现。

一片森林的生态价值不会完整出现在木材收入里。

照护一个家庭的劳动不会完整出现在市场工资里。

一条让人愤怒的内容可能拥有很高点击率，却未必让公共讨论变得更好。

因此，本章不是反对 GDP、利润、分数或点击率。

它反对把任何一个代理指标抬成对现实的全部解释。

AI 使这个问题更紧迫，是因为自动化可以更快执行、更快复制、更快优化。

如果目标设得窄，能力越强，偏差也可能被放得越大。

所以高能力时代真正需要升级的，不只是模型。

还包括目标、权利边界、反馈和纠错。

3.2 麦克纳马拉谬误：当“尸体”成为 KPI

要理解为什么单一维度的“结果导向”容易失真，我们可以回顾越战中“杀敌数”等指标造成的测量异化：它说明代理指标一旦替代真实目的，就可能反过来扭曲行动。

罗伯特麦克纳马拉拥有强烈的数量分析与管理背景。越战时期，美国国防体系在他的领导下更系统地使用伤亡、出动架次、控制区域等量化指标评估进展，“杀敌数”成为其中最具争议的代理指标之一。

问题不在于战争不能统计，而在于一个复杂政治目标被过度压缩成了可以上报的数字。

消耗战逻辑曾把敌方伤亡与补充能力视为判断战局的重要依据；可政治合法性、民众态度、盟友稳定、战略适应与战争目的，并不会因为伤亡表格变漂亮就自动解决。

指标仍然有用。

危险发生在代理指标开始支配行动：越战中的“杀敌数”长期受到夸大、分类不可靠和激励扭曲等批评，数字一旦与晋升、战果和政策叙事绑定，就更容易把“可以统计的结果”误当成“真正实现的目的”。

美国最终未能实现其在越南设定的政治与军事目标，但不能把这场战争的失败归因于一个指标。

战争结果由战略、政治意志、民众支持、盟友与对手行动、地形、国际环境以及美国内部政治等大量因素共同塑造。

“杀敌数”的教训因此更有限，也更可靠：当一个代理指标占据过高地位，它会遮蔽那些难以计量却决定真实结果的因素，并反过来改变组织行为。

丹尼尔克洛奇后用一著名的四步批括“只相信可量化之物”的陷，批后常被“克拉”：

步骤 逻辑谬误 当代翻版

1 测量那些容易测量的 点击率、论文数、股价、GDP

2 忽略那些不容易测量的 心理健康、生态韧性、代际公平、意义感

3 假定没被测量的东西不重要 “那只是情绪”“那只是文化”

4 假定没被测量的东西不存在“社会信任是虚构的”“幸福感无法定义”

在 AI 时代，这种谬误可以被自动化系统以机器规模重复。

推荐系统若只优化点击、停留或互动，就可能把愤怒、信任与长期心理后果留在目标函数之外。

金融模型若只追逐短期收益，也可能低估尾部风险、流动性和系统关联。

教育评价若只盯就业率或分数，就可能看不见判断、好奇、关系与生命意义。

城市规划若只优化通勤时间，也可能遗漏社区连续性、可达性与生活体验。

问题不是“数字会毁掉文明”，而是：当可测量之物开始替代真实目的，能力越强，偏离就可能被放大得越快。

3.3 注意力经济的惨剧：算法如何制造系统性撕裂

如果说越战提醒我们单一指标如何遮蔽真实，那么 21 世纪社交媒体危机则把另一种风险推到眼前：当平台长期优化参与度、传播与停留时间，极端内容、仇恨与误导信息可能获得结构性放大。这里不能把复杂暴力归因于一个算法，但也不能假装推荐机制与传播结构毫无责任。

2018 年，联合国缅甸独立国际事实调查团在报告与公开说明中严厉批评社交媒体，特别是 Facebook，被用于传播仇恨、煽动暴力和制造敌意。它支持的是“平台传播结构可能放大现实伤害”这一判断，而不是“某个推荐算法单独造成了种族暴力”。

发生了什么？难道工程师在代码里写了“煽动仇恨”的指令吗？

并没有。

工程目标常被写成提高参与、相关性、停留或互动；危险发生在这些代理指标压过安全、语境与现实后果时。

一条理性、温和的讨论未必比煽动性内容更容易获得传播。愤怒、恐惧、身份冲突等情绪往往能推动互动；如果系统只把互动当成成功，就可能持续奖励最能触发反应的内容。

模型能够识别模式、预测互动，却不会因为“参与度提高”这一目标本身就自动拥有对真相、尊严与群体安全的完整判断。

于是，“高参与度=任务完成”可能成为危险的代理目标。

在缅甸等案例中，低数字素养、政治与族群冲突、线下权力结构、煽动者行为与平台传播相互作用。把责任全部推给算法会失真；把算法完全排除在因果链之外，同样失真。

真实伤害最终发生在人身上：仇恨被扩大，暴力被组织，数十万人被迫逃离家园。平台不是唯一原因，却已经成为必须承担治理责任的传播基础设施之一。

类似风险并不限于一个国家。

斯里兰卡、埃塞俄比亚等地都出现过社交媒体被用于传播仇恨、煽动和冲突叙事的争议与研究。不同事件的因果链并不相同，不能编造统一“参与度倍数”或宣称算法单独点燃冲突；真正可比的是：当高情绪内容传播更快时，平台是否拥有足够的语境理解、风险响应与问责机制。

这让问题从“某个算法是否邪恶”回到更准确的位置：平台选择优化什么、忽略什么，出了错由谁发现、谁能停止、谁承担后果。

2024 年前后，全球多场选举把生成式深伪和合成内容推入选举安全议程。并非每一次深伪都改变结果，但一旦伪造内容可以在核验之前完成传播，算法仍会面对同一个诱惑：什么最能留住注意力，就优先放大什么。

这是一个典型的“局部最优，全局崩溃”。

维度 局部（算法视角） 全局（文明视角）

缅甸 日活用户创新高，广告收入增长 社会撕裂，族群对立，70 万人沦为难民

斯里兰卡 用户时长飙升 教堂清真寺被焚毁，宗教和解倒退数十年

埃塞俄比亚 内容分享量破纪录 内战升级，人道主义危机

全球 数字广告市值万亿美元 公共事实空间解体，民主制度承压

旧文明算法的致命缺陷，在于它是一个“开环系统”。

它只计算系统内部的收益——流量、金钱、股价——却将产生的废料直接排放到系统外部——社会撕裂、认知极化、心理健康危机、民主制度衰败。

在工业时代，工厂排放的是黑烟。

在智能时代，算法排放的是精神毒素。

▣ AI 的能力越来越强，这种“排放”的速度将超过社会自净能力的极限。

我们不能再允许这种“不计算外部性”的算法，继续掌控人类的神经系统。

3.4 古德哈特定律的诅咒：指标的异化

为了构建新文明的算法，我们还必须翻越一座理论大山。

1975 年，英国经济学家查尔斯·古德哈特提出了一著名的格言：

“当一个指标成为目标，它就不再是一个好的指标。”

这就是古德哈特定律。它揭示了人类社会一个深刻的悖论：我们测量世界，世界却因为被测量而改变形态，使测量失效。

苏联计划经济时代，这个诅咒制造了无数黑色幽默：

指标 工厂的应对 实际结果

钉子数量（个） 生产数百万枚极小的钉子 无法使用

钉子重量（吨） 生产一枚一吨重的巨型钉子 无法使用

布匹长度（米） 生产超薄透光的布料 质量崩溃

公交车维修率 干脆不开工，避免维修 市民无车可坐

在当代教育和科研中，这个诅咒依然生效，且没有国界：

·当“论文引用率”成为学术评价的唯一指标，原本旨在探索真理的科学家，被迫变成了“引用互刷联盟”的成员。学术垃圾泛滥，真正的范式突破反而被淹没。

·当“大学排名”成为高等教育政策的目标，校长们把经费从图书馆转向公关公司，从基础学科转向“易发表学科”。排名上升了，教育质量未必。

·当“考试分数”成为学生成功的唯一标准，教育异化为“提分训练”。好奇心死了，创造力枯了，人赢了考试，却输了人生。

这就是“代理指标”的局限性。

我们想要的是“繁荣”，但因为繁荣很难测量，我们用 GDP 做代理。

我们想要的是“知识”，但因为知识很难测量，我们用分数做代理。

我们想要的是“信任”，但因为信任很难测量，我们用粉丝数做代理。

我们想要的是“健康”，但因为健康很难测量，我们用门诊量做代理。

AI 极度擅长“针对代理指标进行过拟合”。

给它“分数”作为目标，它会找到一万种作弊的方法来提高分数——猜题、刷题、背标准答案——而完全忽略学习的本质。

给它“点击率”作为目标，它会用尽一切心理学操纵手段提升用户时长——愤怒、恐惧、色情、阴谋论——而完全忽略人的幸福。

给它“利润”作为目标，它会将成本外部化到环境、社区、后代身上——而完全忽略可持续发展的真义。

这不是 AI 的错，这是我们的错。

我们给了它错误的目标函数。

面向未来的文明治理不应再用另一个单一、静态、可被博弈的“标量”替代旧指标。

它至少需要从单一标量转向多维判断：不能只问“得到了多少”，还要问“目的是什么”“代价由谁承担”“权利是否被越过”“系统风险是否积累”。

3.5 失序的诱惑：为什么熵不能替文明决定方向

那么，什么才是正确的方向？

物理学家埃尔温薛定谔在 1944 年的《生命是什么》中，借热力学语言讨论生命如何维持高度组织状态。这一思想很有启发，但也最容易被社会理论滥用。

热力学第二定律在适用条件下约束能量与熵的变化；对孤立系统，更准确的说法是总熵不减少。“宇宙热寂”则是建立在特定宇宙学假设上的远期推演，不能被直接搬来证明社会制度。

生命之所以令人惊叹，不是因为它“战胜热力学”，而是因为生命作为开放系统，通过持续交换能量与物质维持、更新并生成复杂结构。

薛定谔曾使用“负熵”一类表达启发人们理解生命秩序；用现代语言说，生物体依靠代谢与环境交换维持远离平衡的结构。把它写成“生命从宇宙榨取秩序”，可以是文学，却不能当作精确的社会科学公式。

文明可以借这一类比提醒自己关注失序、耗散与修复，但社会中的权利、信任、意义、责任和合作并不是热力学变量。

然而，仅靠“负熵”维持有序，最多帮助我们描述某些失序风险。文明跃迁真正改变的，不是把“利己”切换成“利他”这一个开关，而是改变利己实现自身的默认路径：让越来越多的自我成就，不再以损害他人与共同未来为代价。

文明真正做的，是在不确定、冲突与有限资源中建立制度、关系与共同生活，使人能够存续、幸福，并有空间生成意义。

因此，新文明算法不能把 GDP、所谓“社会熵”或任何单一变量当作文明方向本身。真正需要看见的，是现实影响、信任、连接、责任与外部性如何变化，以及成就别人是否也越来越能够成就自己。

新文明算法不能只看 GDP、点击率、利润率这类单一指标，也不能用另一套总分替代它们。信任、连接、真实贡献、责任承担与系统风险可以分别显影；它们帮助判断，却不被压成统一效用。

让我们重新审视前文的案例：

行为 旧算法评价 失序预警

社交平台放大极端内容 参与度↑ - 局部指标成功 信任与安全风险↑ - 需要重新判断

麦克纳马拉的“杀敌数”KPI 尸体↑ - 成功 仇恨↑、战略混乱↑ - 熵↑ - 失败

工厂生产巨型钉子 吨位↑ - 成功 资源浪费↑ - 熵↑ - 失败

高校追逐论文引用 排名↑ - 成功 学术泡沫↑ - 熵↑ - 失败

反过来，那些被旧算法完全忽略的行为，在熵视角下却具有极高的价值：

案例：维基百科

成千上万的志愿者在没有传统工资关系的情况下，投入大量时间整理、考证、修正、维护人类知识。

从旧算法看：这是“浪费时间”。

从熵视角看：这是将杂乱无章的信息——高熵——整理成有序、可信的知识网络——低熵。

维基百科是一个极有力量的例子：大量维护、协作与纠错，可以把分散信息变成更可用的公共知识。

案例：孟加拉乡村银行

格莱珉银行等小额信贷实践曾尝试让传统金融体系之外的人获得信贷机会。其效果、成本与适用性长期存在研究与争论；它不是一种自动把“绝望变成希望”的单向机制。

从文明跃迁的角度，它至少提醒我们：制度设计可以改变谁有机会进入经济关系，但金融工具不能替代权利、公共服务与更广泛的发展条件。

这不是“社会减熵”的证明，而是关于接入、信任与机会结构的现实参照。

案例：原住民的森林守护者

许多原住民共同体长期生活高生物多样性地区，并拥有重要的生态知识与土地治理实践。不同共同体的历史、制度和现实处境差异很大；这里关注的是长期生态维护的价值，而不是把它们塑造成同一种“森林守护者”，评价其价值也不能只看 GDP 是否增加。

从文明跃迁看，真正值得关注的是：长期维护生态、知识与共同资源的工作，如何不被只奖励短期开发的制度持续忽略。

有时，没有把一片土地迅速变成商品，本身就是一种需要被看见的长期责任。

这给了我们一把打开未来的钥匙。

未来高能力 AI 不应被单一点击率、利润或其他代理指标无条件支配。

更稳健的方向是：在明确目的、基本权利、授权、责任和安全边界下进行多目标分析，并把系统失序、连接断裂与长期外部性作为重要风险信号之一，而不是再建立一个“最大化有序度与连接度”的单一总目标。

这不是从物理学直接推出的社会公理，而是一种借助系统视角形成的治理启发。

它不依赖于任何特定的宗教、文化或意识形态。

它是否成立，最终要由现实社会中的权利、制度、行为后果与长期实践检验，而不能由热力学类比代替证明。

3.6 网络结构的启发：如何看见连接与撕裂

如果我们不再只用 GDP 和点击率作为指挥棒，而希望同时看见连接断裂、信息隔绝、冲突扩散与信任受损，那么网络科学能够提供一部分观察工具。

但它首先只能回答结构问题：谁与谁相连、信息怎样传播、网络是否过度集中、哪些节点或桥梁正在消失。

它不能直接计算“良心”。

也不能把“信任”压缩成一条边。

更不能从一张图自动推出谁善谁恶。

可以测量的是某些结构信号；必须由现实证据、语境和人的判断解释这些信号意味着什么。

网络科学确实拥有度数、中心性、模块度、代数连通度等指标。它们对定义明确的图结构有数学意义，但结论高度依赖节点、边、方向、权重和时间窗口如何定义。

想象两个人的网络行为：

人物 A——“流量捕手”

他制造了一条极端谣言，短时间获得大量转发。表面连接数上升，并不意味着社会关系更健康：传播可能同时强化同温层、敌意与跨群体误解。

原本可以对话的群体也许因此减少真实接触，开始互相拉黑、彼此谩骂。

某些网络指标可能捕捉到社群分化、桥接关系减少或传播集中，但不存在一个天然叫做“恶性模块度”的数学结论。

即使网络看起来更“连通”，如果连接主要由仇恨、胁迫或虚假信息构成，也不能因此判定它更有文明价值。

这不是“拓扑学意义上的熵增”，而是需要结合内容、行为后果和真实关系理解的社会失序。

人物 B——“社区织工”

他调解了一场邻里纠纷。从拓扑图上看，他修补了两个断裂节点之间的连接。

他让两个冷战多年的邻居重新说话。

他让互不信任的警民之间有了第一次对话。

对一张定义明确的无向图，Fiedler 值可以反映代数连通性；但连通性高低本身没有道德方向。

一个犯罪网络也可能高度连通，一个受压迫群体为了安全保持低连接也未必是坏事。

所以，网络指标能帮助我们发现结构变化，不能替文明作价值判断。

AI 也不需要被放到“数学上帝”的位置。

我们需要的是把 Fiedler 值、模块度、传播速度等当作观察仪器，而不是当作良心。

行为/情景 可观察的网络信号 仍需回答的问题

煽动对立、制造仇恨、信息茧房 可能出现社群分化、桥接关系减少、传播集中 真实伤害是什么、因果链是否成立

促进对话、修复信任、跨社群连接 可能出现桥接关系增加、沟通路径恢复 关系是否真实、是否自愿、是否持续

图论的价值正在这里：它让一部分“连接如何变化”变得可见；文明判断仍要回到人、权利、事实、责任和真实影响，以及“让成就别人，也成为成就自己的方式”。

AI 可以处理大规模网络数据，帮助追踪信息、资金或政策影响的传播路径，但前提是数据取得合法、模型假设透明、分析结果可质疑。

它可以提示某个结构正在集中、断裂、扩散或失去韧性。

它不能仅凭一个矩阵宣布谁是“癌细胞”、谁是“白细胞”。

这一视角并非西方学术的专利。

不同文明都发展过市场监督、度量衡、公共责任与社区互助制度。伊斯兰法与城市治理传统中的 hisba 就是其中之一，它涉及市场秩序、公共道德与监督实践；具体职责随时代与地区变化，不能用一段未经核对的古代“名言”替代复杂历史。

真正值得保留的，不是一句漂亮引语，而是一个朴素事实：市场从来不只靠价格运行，也依赖规则、信誉、责任与可被追问的行为。

在印度喀拉拉等参与式地方规划、民众科学与社区资源调查实践中，居民知识曾被用于理解本地资源与公共问题。它们与图论指标不是同一种方法，不能强行说成“数学直觉完全一致”；二者能对话的，只是都在提醒我们观察关系，而不只观察存量。

在冲突后的和平建设中，人们也会观察跨群体接触、共同市场、学校、社区活动与日常关系恢复。不存在一套可以跨地区通用的“信任指数”自动证明和平。

停战协议可以结束某些暴力，却不能替代日常关系的重新建立；同样，一次共同活动也不能自动证明长期和解。

新文明的“可观察性”真正应学习的，是这种谦逊：让关系、后果与真实生活更可见，但不把复杂社会压缩成一个万能分数。

3.7 ESG 评分的困境：当代理指标反客为主

在建立新算法之前，我们必须审视一类反复出现的治理困难：当道德目标被转换成评分，评分本身会不会逐渐取代真实世界的结果？

ESG 评分体系——环境、社会、公司治理。

ESG 本意是好的。它试图引导资本向善，让投资者在选择标的时，不仅看利润，也看企业的环保表现、社会责任、治理水平。

问题不在 ESG 三个字，而在任何代理指标都可能遭遇古德哈特定律：当一个分数开始决定资本、声誉与资格，被评分者就会学习如何优化分数。

2022 年 S&P 500 ESG 指数年度再平衡提供了一个很有启发性的案例。

特斯拉被移出该指数。

与此同时，埃克森美孚仍然是该指数自 2019 年推出以来持续在列的大型公司之一。

把这一结果直接解释成“石油公司比电动车公司更环保”是错误的。

这个指数回答的是一套具体方法学问题：在排除规则之后，依据 S&P DJI ESG 分数和行业内相对排序选择成分股；它并不是一张跨行业的“谁对地球贡献最大”排行榜。

因此，特斯拉被移出不能简化成“因为不会填表”，埃克森美孚仍在其中也不能证明其现实气候影响更优。

真正值得追问的是：当方法学、披露、争议事件、行业相对排名和可获得数据共同进入一个分数，使用者是否还记得——分数只是代理变量。

马斯克曾公开把 ESG 称为“骗局”；这可以作为争议的声音，却不能替代对指数方法和真实影响的分析。

所以，这个案例不是“ESG 已经被证明无效”。它证明的是另一件更普遍的事：任何道德评分，一旦被当成现实本身，就会制造认知错位。

我们真正需要衡量的不是“表填得多漂亮”，而是行为造成了什么后果；但后果也不能被一个单一数字穷尽。

另一个更直接的案例，是 DWS 的 ESG 披露与投资流程争议。

其前可持续发展负责人 Desiree Fixler 曾公疑公司 ESG 整合程度的表述，引发监管与市场关注。这些指控最初属于举报与争议，不能在未经裁判时直接写成“全部捏造”。

但后续监管行动确认了实质问题：2023 年，美国证券交易委员会认定 DWS Investment Management Americas 就其 ESG 投资流程作出重大误导性陈述，并存在相关政策与程序实施缺陷；DWS 就 ESG 事项同意支付 1900 万美元民事罚款。

这仍然不证明“ESG 天然虚假”。它证明：只要一个标签能够带来资金与声誉，标签背后的流程、数据与证据就必须能够被核验。

否则，责任投资可能退化为责任叙事。

指标越有力量，验证指标的制度就越重要。

这两类案例给新文明算法留下的教训，不是取消衡量，而是拒绝把衡量结果神圣化：

教训 含义

代理指标替代 把披露、评分或标签误当成真实世界结果

证据不足 高风险主张缺少独立、可复核的事实支持

激励错配 被评价者、评级方法、数据供应与资本使用之间都可能存在利益与方法偏差

复杂度幻觉 以为一套全球统一评分可以穷尽不同产业、法域与社会关系中的伦理问题

如果不解决证据来源、方法透明与可申诉问题，任何道德打分系统都可能演变成新的官僚游戏。

如果不解决“指标与被指标”的博弈问题，任何量化体系都会被过拟合和操纵。

因此，后文讨论溯源、隐私保护证明与责任主体，不是为了制造一种“永不撒谎的数据”。

自述仍然重要，但在高风险场景中不能成为唯一证据；外部记录、独立核验、反方材料、现场事实与申诉同样重要。

我们需要的是可验证、可追溯、能够被质疑和纠正的证据链，而不是把任何数据库、区块链或传感器数据宣布为“不可篡改的现实”。

3.8 算法的普遍困境：三个思想情景

ESG 的争议揭示了一个更普遍的结构问题：任何自上而下、单一维度、脱离语境的量化评价，都可能把指标当成目标。下面用三个明确标注的思想情景检验同一逻辑。

情景一：贫困指标与资源分配

设想公共资源优先流向某个“最贫困”分类，而一旦社区跨过分类线，支持立即大幅退出。

算法逻辑：越贫困，获得的支持越多。

如果退出机制设计不当，指标就可能制造新的激励扭曲：改善处境反而意味着突然失去支持。

问题不是“扶贫制造贫困”，而是任何分档制度都必须认真设计边际退出、过渡与真实改善的激励。

情景二：教师出勤与教育质量

设想学校只把摄像头记录的“人在教室”与工资或考核挂钩。

算法逻辑：出勤可见－出勤提高－教育质量自然提高。

教师可能更准时出现，却仍不能由此推出学生真正学会了什么。

指标可以测到出勤，目标却是学习；二者相关，但不是同一个东西。

指标成功，不等于目标成功。

情景三：反欺诈模型与边缘用户

设想支付平台把“偏离主流交易模式”高度关联为欺诈风险。

算法逻辑：异常模式 = 高风险。

它可能提高筛查效率，也可能误伤那些因为地区、职业、移民身份或非标准金融习惯而呈现不同模式的合法用户。

如果没有解释、复核和申诉，风险识别就可能变成对“不像多数人”的自动惩罚。

这些案例的共同教训是：

任何不扎根于本土认知、不尊重既有社会结构、不设计反馈和申诉机制的算法系统，最终都会产生严重的“脱离效应”。

它不是恶意的，但它无意中执行了文化暴力。

如果算法进入公共或高后果事项，至少应从一开始具备这些受约束接口：

原则 含义

可争议性 任何算法决策都必须允许人类申诉

可解释性 任何拒绝都必须给出因果链，而非黑箱概率

可退出权 社区或个人有权选择“不被算法治理”

本土化接口 全球统一的指标必须允许本地权重调整

3.9 算法进入公共事务前的三条边界

基于前面的结构性教训——不能只看单一指标；

基于 ESG 等现实争议——不能把自述或标签当成结果；

基于三个思想情景——不能把指标与语境、权利和申诉分开；

我们可以提出新文明算法需要守住的三条上位边界。

这些不是由本书直接生效的“全球宪法”，而是思想层面的约束：无论算力多强、模型多先进，都不能让能力自动越过权利、授权、责任与纠错。

边界一：基本权利不能成为效率函数的奖品

任何涉及基本生存条件的自动化系统，都不应仅凭效率或评分把人推到生命与基本权利底线以下；具体保障标准与法律效力必须由现实法域、主体与制度确定。

自然选择的生物学描述不能直接成为人类制度规范；“淘汰低效个体”更不能被算法借作公共权利规则。

在人类文明中，人的基本权利不是效率函数的奖品。

无论一个人处于何种能力、就业或社会参与状态，算法都不能仅凭效率判断剥夺其基本权利。任何合法限制都必须来自有权主体、适用规则、事实、正当程序与责任链，算法只能在被授权范围内执行。

这既是人本底线，也是防止目标函数把个体当成可牺牲变量的制度边界。

系统不能以“大局最优”为名自动取消某个人的基本权利。

一旦“总体收益”被当成可以无程序地牺牲个体的许可证，任何人都可能在下一次优化中成为代价。

技术系统可以把基础生存底线、停止条件与人工接管接口实现为不可绕过的安全约束，但这些约束本身不创造权利来源，也不能替代现实法律、治理授权与救济程序。

边界二：价值与规则必须有正当来源，并且允许修订

定义“什么是有价值的”不能由单一政府、公司、AI 或任何单一群体永久垄断；但这也不意味着存在一个对所有事项通用的统一投票主体。

价值权重、指标和规则的形成与修改，必须回到具体事项的有权主体、适用制度、受影响关系、专业程序与明确授权；贡献记录本身不产生参数投票权或公共决定权。

谁来决定“环保”比“经济”重要？

谁来决定“照顾老人”比“写代码”贡献大？

谁来决定“本土文化保护”应该占多少权重？

在不同现实制度中，市场、公共规则、组织治理、专业判断与社会参与分别承担不同功能，不能被压成同一种“权重来源”。

新文明需要的是可争议、可说明、可修订的正当程序：不同领域可以采用选举、代表、协商、专业程序、利益相关参与或其他合法机制，但不存在一套由贡献值统一加权的普遍民主公式。

拥有贡献记录，只说明真实贡献曾经发生；它可以在相关事项中提供经验、责任或事实信息，却不能因此获得比没有贡献记录者更多的一般政治权利。

算法不是上帝，也不是“公意”的数字化主体；它只能表达和执行已经通过正当程序形成、且能够被审计和修订的规则。

当事实、权利、目的或规则发生变化时，算法及其参数必须随之重新判断、重新授权或停止执行。

真正防止“价值锁定”的制度保险，是保留争议、复核、退出、修订、停止和重新授权，而不是把价值决定转移给一种新的加权算法。

边界三：高后果计算必须可说明、可复核、可申诉

无法说明、无法复核、无法申诉的高后果自动化决定，不得成为涉及基本权利或重大公共利益事项的唯一最终依据。

在涉及人类权益的事项中，AI 如参与分析，至少应能够说明事实依据、规则来源、关键推断、不确定性与可供复核的记录；解释义务不能替代有权主体的决定责任。

例如移民许可、贷款、信用、公共福利或其他高后果事项，AI 可以筛查材料、发现异常、形成建议，但是否产生具有约束力的拒绝、限制或资格后果，必须取决于相应法域和制度中的有权主体、程序与授权，而不是由模型输出自动生成。

可说明的输出应至少区分：

“已确认事实／仍待核验事实；适用规则及其来源；模型或专业推断；不确定性；建议结果；由谁有权作出最终决定；可用的复核、申诉与纠错路径。”

申诉、复核与救济来自制度和权利本身，不是 AI 在作出决定后再额外“赋予”的许可。

如果 AI 的推断有误——例如把行为艺术误判为恶意破坏——错误必须能够在生效前或生效后被人类复核、纠正、暂停，并反向更新模型、规则或使用边界。

黑箱模型在高后果公共事项中的使用必须受到与风险相匹配的限制；具体法律效力由现实法域决定，本书不把单一技术原则直接写成全球统一法律结论。

这就是新文明算法需要守住的“宪制边界”：算法能力越强，权利来源、授权链、责任链和纠错接口越要清楚。

它不规定具体的经济政策，不预设特定的文化价值，不绑定于任何意识形态。

任何算法若进入公共或高后果场景，都必须先通过这些边界检验；通过检验仍不等于自动取得治理权。

定律 保护的对象 防止的风险

基础存在不可否定 每一个体的生存尊严 系统性清算、效率暴政

正当来源与可修订性 共同体的价值与权利边界 单一主体锁定、算法越权

可说明、可复核、可申诉 个体的认知与程序权利 黑箱支配、责任失联

3.10 救生艇：当算法不能替社会给生命排名

有六个人落水，救生艇只能先救五个。

其中有老人，有孩子，有孕妇，有曾经犯过罪的人，也有成就卓著的人。

这一刻最危险的捷径，是让系统偷偷把“谁更有社会价值”变成排序依据。

高后果分配并不需要一套新的“文明对齐测试”。它需要更朴素的边界。

第一，生命与基本人格地位不因财富、职业、贡献、家族或声望被算法重新定价。

第二，年龄、伤情、失温风险、孕产状态、即时可救性等事实，只有在既有法律、医疗规范或正当形成的救援规则明确相关时，才能进入专业判断。

第三，如果规则仍无法区分，就使用事先公开的公平程序，而不是在事故现场临时发明生命等级。

AI 可以帮助计算路线、时序和生存概率，也可以执行预先授权的安全动作，尽量扩大所有人的获救机会。

但它不能因为算得更快，就取得“谁更值得活”的价值主权。

真正需要被检验的，是系统有没有分清：哪些是事实，哪些是专业判断，哪些是有权主体已经形成的规则，哪些根本不该被拿来排序生命。

艰难选择不会因此消失。

文明的最低要求，是别在艰难时刻顺手把人变成一张社会价值排行榜。

3.11 技术形态会变，治理边界不能一起消失

未来计算架构可能继续变化，量子计算也可能在某些任务上改变速度、密码安全或系统设计。

^[8]

今天没有必要为了这些尚未确定的组合，提前写出一整套“量子-AGI 伦理学”。

真正稳定的问题与技术名称无关。

系统越强、执行越快、内部越复杂，就越要在系统之外保留足够的证据、约束和停止能力。

我们未必能够复原一个高能力系统内部每一步推理。

但必须能够知道它用了什么权限、接收了什么关键输入、产生了什么高后果动作，以及谁有权停下它。

技术形态可以变化。四条边界不能因此消失：

关键权限有来源。

高后果动作可停止或接管。

运行留下足够审计证据。

现实责任仍然有人承担。

量子、经典或未来其他计算方式，都不能仅凭能力差异自动改写这四条关系。

3.12 本章结论：算得更值得

我们从波斯特洛姆的“回形针”噩梦出发，

穿越了麦克纳马拉的尸体统计、缅甸的仇恨算法、ESG 的伪善表格、墨西哥的贫困陷阱、印度的摄像头教室、肯尼亚的被误读的哈瓦拉，

最终站在了新文明算法的大门口。

我们得出的结论是：

我们不能停止计算。

在高度互联、AI 能力持续增长的复杂系统中，仅靠直觉和道德说教不足以处理全部复杂性。我们需要算力、算法与模型辅助理解和预演风险，但这些能力不能自动取得决定权。

但我们必须改变计算的目标。

旧文明的计算公式是：

Maximize(资本回报率)

Maximize(GDP 增长率)

Maximize(点击停留时长)

Maximize(论文发表数)

目标函数错了。算力越强，毁灭越快。

新文明的计算需要同时守住：

降低系统性毁灭概率

维护真实连接与共同生活条件

让局部理性不再持续制造整体失序

新文明的计算，不再把资本回报、GDP、点击停留时长当作唯一目标；它要帮助看见长期后果、真实连接与共同风险，让合作、维护、照护和创造不再在计算里自动消失。

这不是从“追求增长”到“放弃增长”的退步。

这是把增长重新放回文明目的之下。

增长可以是手段，存续必须守住底线。

效率可以是工具，不能成为价值的主人。

局部可以追求自身利益，但不能把共同世界当作无限外部性。

本章核心提炼：

文明跃迁的根本，不是更努力地对抗熵增，也不是要求人从此“利他为先”。它改变的是利己实现自身的默认路径：让局部获利越来越不必以系统撕裂为代价，让帮助、维护、照护、连接与共生也能得到真实回应。熵增可以解释某些失序，却不能替文明决定价值方向。

“赢家通润”和贡献记录，只能被理解为这种路径变化的制度显影之一：让真实贡献被看见、承认和回应，而不是把贡献自动兑换成基本权利、一般权力或人格等级。

如果未来越来越多技术与制度把真实影响、权利、责任和长期后果纳入设计，激励结构可能逐步变化；但不存在一套“新算法”会自动在每台服务器和芯片中运行，更不能把制度转型写成技术必然。

如果“赢”不再意味着“赚最多的钱”，

那么“赢”意味着什么？

接下来的第四章，我们将探讨这个驱动力的转换：

从竞争到贡献。

从“赢家通吃”的角斗场，

到“赢家通润”的共生体。

第四章 从竞争到贡献：改变相遇的结构

“竞争是为失败者准备的。”

——彼得·蒂尔，PayPal 联合创始人

“生命从来不只靠竞争向前；有些最重要的跃迁，恰恰发生在彼此不再只把对方当作敌人时。”

——本书，据林恩·马古利斯关于共生演化的研究与思想转述

“我之所以成为我，也发生在我们之间。”

——据 Ubuntu 关系性思想转述

真正需要重构的，不是人的私心，而是私心被奖励的那条路。如果压过别人总是最便宜的成功方式，再高尚的劝告也很难改变系统。

4.1 稀缺时代的遗产：为何人类迷恋角斗场

1. 罗马角斗场的余音：竞争的血色底色

古罗马的科洛西姆角斗场里，四周是震耳欲聋的欢呼声，脚下是混着黄沙的鲜血。两名角斗士持剑对峙，只有一个人能活着离开。

对于观众来说，这是一场表演；但对于角斗士来说，这是最纯粹的“零和博弈”：我的生存，必须建立在你的毁灭之上。

现代社会并没有告别冲突，也没有被一种“角斗场逻辑”完全支配。更准确地说，零和竞争在不同制度中反复出现，并会换成排名、份额、资格、流量等新的形式：

在华尔街，它是屏幕上闪烁的红绿指数，是收割对手盘的快感；

在写字楼，它是为了那 5% 的晋升名额而进行的“末位淘汰”；

在教育中，它是为了挤进名校而进行的“一分干掉千人”的惨烈内卷。

我们为什么对此如此迷恋？甚至到了不竞争就感到不安的地步？

2. 竞争为何反复出现：匮乏、风险与制度

我们没有必要用一段简单的“自私基因故事”解释全部竞争。长期的资源约束、不确定性、群体边界与权力结构，使争夺在历史上反复成为生存策略之一；合作、互惠与照护同样古老。

在狩猎采集、农业与早期国家社会中，食物、土地、时间与安全都受现实约束。某些情境里，抢先占有有利于个体；另一些情境里，共享猎物、共同防御、抚育后代与互惠网络反而提高群体存续概率。

文明真正继承的，不是一条“竞争必胜”的生物程序，而是对匮乏、失去与被排除的高度敏感。

所以，利己并不等于必须伤害别人，合作也不等于取消自身利益。问题从来不是消灭利己，而是旧制度是否把“压过别人”塑造成实现自身利益的默认路径。

这种对得失与位置的敏感进入现代社会后，会被工资、排名、晋升、流量、声望和稀缺资格重新编码。

赢会带来奖励、安全与被承认；输可能带来收入下降、机会收缩、羞耻和不确定。

这些反应有生理基础，也受文化、关系和制度塑造；它们不是一套不可修改的文明“出厂设置”。

我们真正需要警惕的是：当技术已经改变部分稀缺结构，制度仍不断把人推回“只有踩住别人才能站稳”的角斗场。

3. “红皇后效应”：跑得越快，原地踏步

在演化生物学与组织研究中，人们常借“红皇后效应（Red Queen Effect）”描述一种相对竞争：你不断加速，只是为了不被甩下。这个名字来自《爱丽丝镜中奇遇》里红皇后那句著名的奔跑隐喻。

工业化和现代市场竞争在一些领域把相对竞争放大到很高强度，但不同国家、行业和制度中的程度并不相同。

当一个外卖平台用更精细的算法不断缩短预计配送时间，单个骑手初期也许能凭更快速度获得更多订单；但一旦更快节奏被写进新的基准，个体优势很快就会变成全体压力。

结果可能是：每个人都跑得更快、更累，收入与地位却未必同步提高，交通安全、休息和身体成本反而上升。

这可能形成一种系统性耗散：参与者投入更多生命时间争夺相对位置，却未必形成与消耗相称的共同增量。风险在于竞争基准不断上移，而不是“所有人已经被拉进同一个角斗场”。

竞争被神圣化了。

19 世纪工业化留下了一个复杂遗产：竞争确实推动过降本、扩产、技术传播与组织创新，也伴随着垄断、劳资冲突、危险劳动和外部成本。铁路、钢铁与现代城市不是由“你死我活”单独创造的，而是资本、技术、国家、劳动、制度和竞争共同作用的结果。

因此，承认竞争曾经释放巨大生产力，不等于把任何竞争都神圣化，更不等于把失败者承受的全部代价视为进步的必要祭品。

从明治日本到独立后的印度，再到资源与腹地有限的新加坡，现代化叙事都曾把国家生存、工业能力与国际竞争紧密相连；但各自历史条件、制度选择和发展路径不同，不能用几句未经核实的名言把它们压成同一种“竞争信仰”。

这些成功经验在不少制度与组织中形成了路径依赖，使“更多竞争=更多进步”容易被当作默认答案。

一些现代制度和社会叙事反复假定：

- 只要引入竞争，效率就会提高
- 只要允许优胜劣汰，文明就会进步
- 排名制造焦虑，但焦虑制造奋斗
- 失败者的眼泪，是胜利者祭坛上的合法祭品

我们将这种逻辑复制到了教育——考试排名、重点学校、精英选拔；

复制到了医疗——市场化改革、三甲医院评级、患者用脚投票；

复制到了科研——影响因子、项目竞标、人才帽子；

复制到了家庭——“别人家的孩子”是最锋利的教鞭。

然而，我们忽略了竞争生效的边界条件。

竞争更可能具有正当性并产生建设性结果，通常需要若干边界条件；下面两项只是其中重要的两项：

假设 含义 失效条件

稀缺性假设 关键资源与能力受现实约束，无法无条件满足所有需求 AI 与自动化降低部分生产与复制成本，并重组稀缺结构，但不消除物质、能源、时间、空间、生态与能力约束

可逆性假设 失败者可以退出、重来、换赛道 算法锁定了赛道，且失败累积为系统性边缘化

当时间来到 21 世纪 20 年代，AI 和自动化的快速发展开始同时改变这两个假设的适用边界：部分数字产品与认知劳动的复制成本显著下降，而算法化机会结构也可能使失败更具累积性。

在数字化和智能化的领域：

- 复制一份既有数字知识的新增传播成本可以极低，但生成、验证、维护、存储、算力与注意力仍有成本
- 复制一套成熟数字诊断模型的新增分发成本可以很低，但数据、验证、设备、临床责任与现实医疗资源仍受约束
- 扩展 AI 劳动力可以显著降低部分认知任务的单位成本，但算力、能源、芯片、数据、维护与现实执行能力并非无限

在部分复制成本显著下降、但其他稀缺仍然存在的领域，如果制度仍机械沿用某些旧的竞争假设，可能出现一种风险：

竞争从促进发现和效率，转为病理性的相对消耗。这里的“癌变”只是警示性比喻，不是社会系统的生物学诊断。

4.2 内卷的数学本质：红皇后与系统性噪音

2020 年代，“内卷”从学术概念进入中文互联网，迅速成为描述过度竞争与低效消耗的高频词。

“内卷”。

“Involution”并非由克利福德格尔茨首创。更早的人类学讨论已使用这一概念；格尔茨在 1963 年的《农业的内卷化》中把它用于分析爪哇农业：在土地、人口与殖民经济压力下，劳动不断密集化，系统内部越来越精细，却没有发生相应的结构跃迁。

在他的分析中，单位土地可以承载更多劳动和人口，但单位劳动者的收益并未随投入同步提高。

人更忙，结构却没有因此打开新的发展通道。

格尔茨让“内卷”成为“增加复杂度却未实现结构转变”的经典社会科学隐喻之一。

半个世纪后，“增加投入却没有打开新通道”的结构，被许多讨论借来描述不同地区的过度竞争；它不是全球劳动者共享的单一经验。

在 AI 介入的平台经济与算法管理中，这种相对竞争可能被进一步加速。

案例：平台劳动中的算法管理

近年的平台劳动研究、劳动者调查与新闻报道反复呈现同一结构：订单分配、时间预测、评分、奖惩和绩效指标越来越多地由软件系统参与。平台、地区与岗位之间差异很大，全球劳动者不存在一份可以由单一报告概括的统一处境。

在仓储与配送场景中，细粒度计时和绩效追踪可以提高可见性，也可能把人的休息、身体差异和异常情况压进同一速度指标。

网约车与零工平台会通过价格、派单、激励、评分和接单规则影响劳动者选择，具体机制随平台和地区变化。

外卖与配送平台常把接单率、预计送达时间、评分与奖惩嵌入骑手日常。人在现实交通、天气与身体限制中工作，软件却可以持续更新目标；这种节奏差本身就可能形成压力。

中国外卖骑手“困在系统里”的讨论并非孤立现象。它与全球平台劳动中关于算法管理、劳动自主性、收入可预测性与申诉机制的争论相互呼应。

这就形成了进化生物学中的“红皇后效应”。

“你必须拼命奔跑，才能保持在原地。”

在《爱丽丝镜中奇遇》里，红皇后对爱丽丝说这句话。

在算法时代，这句话变成了：

“你必须不断加速，才能维持收入不降。”

在这个过程中，竞争新增了什么真实价值，又把什么成本转给了劳动者和社会？

对于消费者，配送可能快了几分钟；对于平台，周转和留存指标可能更好看。

但如果效率提升主要来自不断压缩人的时间缓冲，新增收益与新增风险就必须一起计算。

交通安全、职业健康、收入稳定与社会保障成本，都不能因为没有进入配送时长指标就被当成不存在。

这是“系统性噪音”。

在旧逻辑下，这叫“效率提升”。

如果效率提升主要靠把未计入指标的代价转移给人的身体、时间与安全，这就不是完整意义上的效率提升。

AI 可以放大竞争，也可以降低成本、创造新能力；关键不在技术天然朝哪边走，而在目标、分配与反馈结构把新增能力导向哪里。

如果新增能力只被用来把每个人像干毛巾一样拧得更紧，技术越强，系统消耗人的速度反而越快。

这不再是值得赞美的优胜劣汰。

这是把效率变成自我消耗。

4.3 算法博弈的灾难：当竞争失去摩擦力

如果说配送骑手是肉身的内卷，那么金融市场的算法战争则预演了 AI 时代的另一种竞争灾难——

失控的反馈环。

2010 年 5 月 6 日下午 2:45，美国股市。

道琼斯指数在几分钟内暴跌近 1000 点，近 1 万亿美元市值瞬间蒸发。埃森哲的股价在几秒钟内从 40 美元跌至 1 美分。

联合监管调查没有把这场事件归结为一个单一“真凶”。

当天市场本已受到欧洲债务危机等因素影响而高度紧张；随后，期货市场的一笔大型自动卖出程序进入流动性正在变薄的市场，高频交易者与其他流动性提供者的策略相互作用，冲击迅速跨市场传导。

这不是“算法突然有了恶意”。

后来，英国交易员纳温德辛格萨劳因长期使用欺骗性挂单等操纵手法被追诉；监管机构认为他的行为在 5 月 6 日也加剧了订单失衡。2010 年的联合调查与后续追诉共同指向多因素事件，而不是“萨劳一人制造了闪崩”的单因果故事。

真正值得文明记住的，是市场在高速自动化、流动性骤降和策略相互反馈时暴露出的脆弱性。

高频交易者在事件早期曾吸收部分卖盘；当压力继续扩大时，许多参与者迅速减少风险敞口或撤出流动性，速度本身开始成为反馈放大器。

于是，本来分别合理的止损、撤单、减仓与风险控制，在同一个高速网络里可能相互触发。

步骤 系统行为 可能后果

1 压力市场中出现大额自动卖出 流动性被快速消耗

2 高频与其他交易策略迅速响应 交易量上升、价格冲击被放大

3 部分流动性提供者撤回报价 买卖价差扩大、跨市场传导加剧

4 个别证券一度缺乏正常对手盘 出现极端、随后被取消的异常成交

这就好比在一个拥挤的剧场里，有人喊了一声“着火了”——

然后所有人都试图比别人先跑出去。

结果就是所有人堵死在门口，相互踩踏。

这次闪崩是一个严重的警告：

当自动化智能体高速互动时，个体策略即使各自“理性”，系统也可能收敛到对整体不利、甚至高度脆弱的均衡。

并非必然走向最差结果，但连接越密、速度越快，错误反馈越可能在人工理解和干预之前被放大。

金融市场已经给过这种警告。

2012 年，Knight Capital 的软件部署故障在约 45 分钟内造成约 4.4 亿美元损失，几乎摧毁公司。

它不是因为市场突然“失去理性”，而是错误代码以机器速度执行，直到损失已经扩散。

这类事件说明：当执行速度远快于人类发现、理解与接管的速度，局部错误会在制度反应之前先完成放大。

这些还是发生在被严格监管的金融市场。

如果未来的战争、电网、医疗等关键系统也高度依赖高速自动化协同——

类似的反馈失控机制就可能跨出金融市场，进入更高后果的现实系统。

不同领域的机理与防护当然不同。金融闪崩不是它们的模板，却是一记足够清楚的警钟：速度可以放大能力，也可以先于理解放大错误。

4.4 生物学的另一条线索：林恩·马古利斯与内共生

复杂系统的出路，不能从某一种生物学机制直接抄成社会制度。

但生命史仍提供了一条值得注目的线索：竞争从来不是进化唯一的动力，合作、共生、漂变与环境选择同样参与塑造生命。

达尔文常被通俗叙事压缩成“丛林法则”——仿佛进化只剩弱肉强食。真正的进化论远比这幅图景复杂。

20 世纪，一位曾被主流长期质疑的科学家，把其中一条被低估的线索重新推到人们眼前。

她就是林恩·马古利斯。

1967 年，马古利斯系统推进了真核细胞起源的内共生理论，主张线粒体、质体等细胞结构具有来自古老细菌祖先的演化来源。

她挑战的不是自然选择本身，而是复杂细胞起源叙事中对内共生作用的长期低估。

最醒目的证据，就藏在我们每一个真核细胞里。

线粒体，以及植物和藻类细胞中的质体，保留着清晰的细菌祖先痕迹。

它们拥有自己的遗传物质，并在膜结构、分裂方式和分子系统上呈现出与细菌演化来源相一致的特征。

它们不是一开始就属于今天这套细胞结构。

随后几十年的细胞生物学、分子生物学与基因组证据，使线粒体和质体的内共生起源成为现代生物学的基本认识。

但真正发生在远古的每一步，并没有被完整记录下来。

今天较有把握的是：线粒体祖先来自细菌谱系，并在真核生命早期与宿主发生了极深的长期整合；宿主究竟是谁、最初如何相遇、先后顺序怎样，仍有重要问题处在研究与争论之中。

因此，我们只能依据今天留下的证据重建二十亿年前的过程，而不能把它写成一幕已经被摄像机记录下来的戏。

没有证据允许我们把它具体写成某个“大细菌”在某一天吞下“小细菌”、双方当场达成一份休战契约。

更接近科学事实的图景，是两个原本拥有不同演化历史的生命系统开始共同存在。

它们的最初关系未必天然友好，甚至可能包含利用、冲突与不稳定。

但在漫长的选择、依存、基因转移与功能重组中，边界逐渐被重新书写。

曾经彼此独立的生命单位，最终成为一个更深层次共同系统。

这才是内共生最值得凝视的地方。

不是“敌人一定会变成朋友”。

而是：演化并不只有淘汰这一种结构变化。

有时，真正深刻的变化，是原本分立的生命开始形成稳定而不可轻易拆开的相互依存。

整合之前 整合之后

独立的演化谱系 高度整合的宿主—内共生体系

相对分立的代谢 长期耦合的代谢与能量关系

各自独立的遗传命运 细胞核与细胞器之间持续的协同演化

关系可以松散或不稳定 通过遗传与细胞机制实现稳定传递

彼此边界清晰 边界被重新组织，形成新的生命层级

线粒体的起源，是生命史上一场真正重大的演化转变。它参与塑造了真核生命新的能量、代谢与细胞结构条件。

如果没有这场转变，生命史会走向哪里，没有人能够从今天反推。科学可以确认已经留下证据的路径，却不能把未发生的另一个地球写成事实。

因此，马古利斯真正带来的思想冲击，并不是“共生必胜”。

竞争可以筛选，合作可以整合；生命还会漂变、迁移、分化，也会和环境压力中不断改写自己的边界。

结构性的创新有时来自竞争，有时来自合作，有时恰恰来自原本分立者建立起新的长期依存。

生物学不会替文明写好制度答案。

但它至少打破了一种贫乏的想象：利己与合作并非天然只能站在彼此对面。

当合作能够改变彼此获得利益的路径，成就对方就可能同时成为成就自己的方式。

这种对“关系”的凝视，也并非只存在于现代生物学。

佛常以“因陀”的意象表相互映照、相互入的世界：一珠子的光，入一珠子；一处变化，也在关系之网中留下回声。

苏菲传统中也长期存在关于统一、关系与彼此映照的思想意象。

许多原住民文化则以各自不同的语言强调人与土地、生命、祖先和共同体之间的互惠责任。

这些传统彼此差异巨大，不能被压缩成同一种哲学，更不能被拿来替现代制度背书。

共生属于生命史；关系智慧出现在许多文明经验之中。

它们能够照亮思想，却不能代替证据。真正的文明选择，仍要在人类自己的现实、权利、责任与经验中接受检验。

4.5 文明的新隐喻：从角斗士到线粒体

把内共生等生命史材料作为有限类比，可以帮助我们看到一种可能：竞争之外还存在长期协作与共同受益；生物学本身不替 AI 时代给出制度答案。

某些工业与现代竞争结构可以用“角斗场”作文学类比，但工业时代的人类从来同时存在家庭、合作、公共制度、互助与市场竞争。

我们像细菌一样相互吞噬，争夺有限的市场份额、有限的自然资源、有限的注意力。

线粒体提供了另一幅图景。

不是因为人类必须变成 AI 的“线粒体”，也不是因为 AI 天然就是文明的宿主。

更强的智能，不因更强就获得支配文明的资格。

人类也不需要所有能力上与 AI 竞速，才证明自身价值。

这个隐喻真正保留的，是能力互补、长期协作与共同受益的可能。

意义、关怀、价值判断、关系与在场，仍由真实主体在生活中生成；它们不是为了给 AI 一块“生态位”。

这只是“从竞争到贡献”的一种生命史启发，不是社会制度的生物学证明。

维度 竞争逻辑 贡献逻辑

目标 打败对手，占有资源 服务宿主，成为不可或缺

关系 零和博弈 正和共生

时间视野 短期套利 长期存续

失败代价 被淘汰 被忽视

成功标志 赢家通吃 赢家通润

这种方向不只是一种道德呼吁。

在重复互动、长期依赖和共同风险较强的条件下，它也可能成为更符合长期自利的现实策略；是否如此，需要由具体制度与经验检验。

如果合作不只是道德选择，它能不能也成为一种更好的利己？

4.6 阿克塞尔罗德的竞赛：当利己走向合作

如果人只和世界相遇一次，背叛很有诱惑。

拿走这一轮，下一轮不必负责。

可真实生活很少只有一轮。

我们会再见。

邻居还会住在隔壁。

同事明天还会坐在同一张桌旁。

今天违约的贸易伙伴，明年也许还要重新谈判。

当明天仍会相遇，今天的选择就有了回声。

1980年，罗伯特·阿克塞尔罗德把这个问题交给了计算机。

他邀请不同学科的研究者提交策略，让它们在迭代囚徒困境中反复相遇。

问题不是：谁看起来更高尚。

而是：在反复相遇的世界里，什么样的策略能走得更远？

结果出人意料。

一个极其简单的策略赢了。

它叫 TIT FOR TAT——报还一报。

第一步，合作。

对方继续合作，它继续合作。

对方背叛，它下一轮回应。

对方重新合作，它也立即回来。

它不首先伤害别人。

也不要求自己永远忍受伤害。

它有善意，也有边界；会回应，也会重新打开合作。

更有意思的是，第二轮参赛者已经知道第一轮发生了什么，可以针对结果重新设计策略。

TIT FOR TAT 再次获胜。

这里的“胜利”，不是数学替善良签下保证书。

换一套规则、对手、信息、噪声与退出条件，结果可能不同。

博弈论不能证明“好人必胜”。

但它让一个古老的误会裂开了一条缝：

自利与合作，并不天然站在彼此对面。

当未来被纳入今天的选择，理性的边界就变了。

一次背叛可以赢下一轮，也可能输掉后面的许多轮。

一次合作也不只是牺牲，它可能在建立下一次合作仍然成立的条件。

阿克塞尔罗德真正给文明的启示，不是“去做好人”。

而是：改变相遇的结构，人的选择就可能改变。

如果一个社会让伤害可以迅速获利、让后果轻易外部化、让人与人只剩一次性交易，合作就会越来越昂贵。

如果真实的合作能够得到真实回应，如果守约、帮助、创造与承担能够留下可信的经验，如果明天仍值得彼此相遇——

合作就不必只靠道德坚持。

它也可以成为实现自身长期利益的一条现实路径。

文明跃迁无须消除利己。

只需改变利己实现自身的默认路径。

让成就别人，也成为成就自己的方式。

4.7 Linux 的启示：集市也能建起基础设施

阿克塞尔罗德回答：为什么合作不一定吃亏。Linux 接下来回答：人为什么愿意主动贡献。

1991 年，林纳斯托瓦兹发布了最初的 Linux 内核。

最初只是一个个人项目，此后却被无数开发者、企业、基金会与社区不断接力。

它后来成为全球数字基础设施的重要组成部分。

这段历史比“开源击败商业”复杂得多。

Linux 的发展既有志愿贡献，也有企业工程师；

既有开放协作，也有商业服务；

既有个人声誉，也有组织投入。

它与商业世界并非永远站在彼此对面。

几十年之后，最值得注意的是谁“战胜”了谁。

而是一个没有单一中央命令的开放项目，能够被长期维护、不断分叉、重新合并，并支撑庞大现实系统。

这说明开放协作可以形成复杂基础设施。

它不证明所有行业都应开源，

也不证明市场会消失。

它只让一种绝对判断失效：复杂成果并不只能由封闭的单一组织长期创造。

真正值得追问的是：为什么那么多人愿意把自己的能力给出去，又为什么企业后来也愿意参与同一个开放生态？

埃里克·雷蒙德用“大教堂与集市”描述两种开发方式。

这个隐喻最有价值的地方，不是把世界分成善恶两边，而是让我们看见分布式协作的另一种组织可能。

在开源社区中，贡献会带来声誉、信任与影响；

但声誉不是主权，贡献也不是人格等级。

维度 封闭研发 开放协作

代码访问 依许可与组织边界开放 可在开源许可下共享与修改

协作结构 组织内部为主 社区、企业、基金会与个人交织

激励 薪资、股权、组织使命 薪资、声誉、学习、兴趣与共同需求并存

创新 集中研发为主 分布式试验与维护并存

这种结构释放了一种重要可能：

一个人可能先为解决自己的问题写出补丁，

再把它交给共同体，

别人继续修改，

最终让一个局部需要变成公共能力。

Linux 当然首先是软件。

但它也为文明留下一个现实隐喻：

成就共同体，并不必然要求放弃自己；

开放协作，也不必等于无报酬、无产权或无边界。

当贡献能够得到真实回应，

“我为什么要给出去”就有了比道德说教更丰富的答案。

4.8 奥斯特罗姆的公共池塘：公地不再悲剧

怀疑论者会问：

“如果没有产权和竞争，不会导致‘公地悲剧’吗？大家都去抢资源，没人维护怎么办？”

1968 年，生态学家加勒特·哈丁在《科学》杂志发表了一篇经典论文，题为《公地的悲剧》。

他设想了一个公共牧场：

- 每个牧民都想多养一头牛
- 多养一头牛，收益归自己
- 过度放牧的成本，由所有牧民分摊

理性的个体行为 – 集体的毁灭结局。

哈丁真正强调的是：在缺乏有效约束时，共享资源可能被个体激励推向过度使用。后来的政策讨论常把出路简化为私有化或国家控制，但那不是唯一可能。

这个“悲剧”隐喻影响深远，也长期塑造了人们对公地的想象。

直到奥斯特罗姆用大量实证研究告诉世界：

共同资源并不只剩“私有化或国有化”两条路。

2009 年，她成为首位获得诺贝尔经济学奖的女性。

不是因为她的理论更精致，而是因为她做了哈丁从未做过的事：

她真的去考察了世界各地的公共资源管理实践。

地点 资源类型 管理方式 结果

瑞士托贝尔 高山牧场 村民自治，轮牧制度 500 年持续利用

日本长野 公有森林 村落协议，采伐配额 300 年未退化

菲律宾 灌溉系统 农民协会，水权协商 西班牙殖民时期延续至今

肯尼亚 旱季水源 部落长老，轮值看守 代代相传

斯特姆：

在部分条件具备的真实共同体中，公共资源可以通过多中心、自主治理等方式得到长期维护，并非只能在“完全私有化”与“单一国家控制”之间二选一。

她总结了八条设计原则，其中最核心的是：

原则 内容 与前三卷的呼应

清晰边界 谁是当前合法使用者、权利边界是什么 现实权利与生效使用关系可验证

集体选择 受影响者参与规则形成 有权主体、利益相关参与与正当程序分层

分级纠正 对违反共同规则的行为依规则逐步处理 证据、比例原则、申诉与修复

低成本仲裁 冲突解决机制本地化、快速 可解释性、申诉与独立复核

奥斯特罗姆的大量经验研究表明：

人类不是只能做自私的短视者。

人类可以在清晰边界与共同规则中成为负责任的长期维护者。

在 AI 与全球网络时代，可以把大气、海洋、数据空间、知识公地等跨境问题类比为不同类型的“公共池塘”；它们的产权、法权、治理尺度和物理属性并不相同，这只是公共资源治理的分析入口。

气候变化等跨境问题说明，仅靠彼此独立的国家竞争不足以解决共同外部性；现实仍同时依赖国家政策、国际合作、市场、技术与地方治理，而且成效并不统一。

其中确实存在公地悲剧式激励：局部收益可能集中、部分成本跨境和跨代扩散；但国家行为并不只有“都想多排放一点”这一种动机，也受发展阶段、能源结构、政策与合作机制影响。

我们需要吸收奥斯特罗姆式多中心治理的核心启示：边界清楚、受影响者参与、规则可执行、冲突可解决，并允许多种治理中心在其权限内协作。

贡献值体系若参与其中，只能记录与资源维护、风险承担、责任履行有关的事实；它不能代替产权、使用权、公共决定、专业程序或制裁授权。

·对公共池塘作出维护贡献的人，其贡献应被记录和承认；是否取得使用权，仍取决于该资源的法律关系、共同规则、合同或明确授权。

·对公共池塘造成损害的人，应依事实、责任与适用规则承担修复、赔偿或必要限制；系统不能仅凭贡献分值自动剥夺权利。

奥斯特罗姆的研究提醒我们：

公共资源并不必然走向悲剧。

长期有效治理通常需要清晰边界、适配规则、受影响者参与、监督、纠正与冲突解决等多种条件；不能压成“使用者共同制定规则+惩罚”两项充分条件。

4.9 稀缺结构的迁移：从单一物质约束到多重瓶颈

从当前事实看，AI 正在改变部分信息、认知任务和数字内容的成本结构；从更长时间看，能源、算力、可信信息、注意力、土地、生态和现实执行能力的相对稀缺性还可能继续变化。这里讨论的是稀缺结构的迁移趋势，不是一个已经完成的“AI 时代经济学定型”。

“稀缺搬家”是一种文学压缩。土地、能源和物质稀缺并没有退出；新增或变得更显眼的瓶颈包括算力、可信信息、注意力、验证能力、生态承载与不可逆选择。

经济学是研究稀缺资源配置的学科。

场景/时期（仅作典型比较） 重要稀缺 常见分配机制 典型权力来源（均非唯一）

部分农业社会 土地、水、劳动力、粮食 继承、习惯、市场、公共规则、武力等 家庭、地主、共同体、国家等

部分工业社会 资本、能源、设备、技能 市场、产权、组织与公共制度 企业、资本、劳动组织、国家等

平台化信息环境 注意力、数据、可信来源 推荐、市场、订阅、公共规则 平台、媒体、用户与监管机构等

当前及未来高智能环境 算力、能源、可信关系、注意力、关键资源与现实能力 市场、权利、验证、公共制度与多元协调 多主体网络与制度；具体形态尚未定型

原子层面的稀缺不会消失，但部分物质生产约束可能被技术显著降低。

聚变能源、先进制造与分子工程如果持续成熟，可能显著降低部分能源与物质生产成本；但可用能源的获取、转换、储运，关键材料、土地、设备、维护、生态承载与安全约束仍然存在。技术改变的是稀缺结构和瓶颈位置，而不是把物质世界变成无约束的无限供给。

比特的复制可以极其廉价，但高质量信息、可信来源、注意力、算力和验证能力仍然稀缺。

AI 可以大规模生成文本、图像、视频、代码、音乐和 3D 模型，使内容供给迅速扩张；与此同时，注意力、可信度、原创判断、算力、数据质量和现实验证反而可能成为更关键的瓶颈。

在深伪、自动生成与大规模信息传播成本下降的环境中，可验证的来源、履约记录与关系信任可能变得更重要；它们仍具有事项、关系与时间边界。

如果合成内容与自动化攻击能力继续扩散，验证现实来源的成本可能继续上升。在这种风险环境中——

“这项事实、身份或承诺在当前事项中是否可信？”

这一判断可能变得越来越稀缺和有价值，但不应被压成可交易的统一“信任资产”。

行为 对信任的影响 在 AI 时代的稀缺性

撒谎、欺诈、造谣 消耗信任 易生产，低成本

诚实、履约、担保 生产信任 难验证，高风险

某些短期、一次性、强对抗的竞争激励可能消耗信任。

当制度奖励短期结果且缺乏核验时，撒谎可能获得局部收益。

长期履约、真实贡献、责任承担和可验证修复可能积累特定关系中的信任。

只有当事实可核验、造假有成本、履约有长期回报时，诚实与负责才更可能成为稳定策略。

因此，AI 时代的分配与协作不能再只看“谁占有了多少”，也不能反过来建立一个以“信任总分”统一排序所有人的新社会。

信任是协作中的重要事实，但它具有关系、事项、时间和情境边界，不能被压成一种普遍购买力。

贡献值的定位是对真实贡献及其后果的受约束记录：它不是货币、不是社会信用总分、不是人格评分，也不是一般资源资格或治理权的自动来源。

这解释了为什么文明需要让贡献被看见——

不是要求所有人通过贡献证明自己更高尚，

而是避免那些真实支撑他人和共同世界的付出在制度中长期不可见，同时不因此创造新的等级货币。

在信任稀缺的场景中，长期履约、真实负责与可核验贡献可以降低特定合作的不确定性；

但任何合作资格、资源使用或公共权利仍应回到相应事项的事实、规则、权利与授权，而不是“谁的社会信用更多”。

4.10 赢家通润：一个关于流体的隐喻

“赢家通吃”是工业、金融和平台竞争中反复出现的一种逻辑，但不是工业文明的唯一座右铭。

胜者拿走一切，输者从零开始。

- 标准石油公司控制了美国 90% 的炼油能力
- 脸书收购 Instagram 和 WhatsApp，消灭潜在竞争对手
- 亚马逊用低利润挤垮书店，再用垄断定价回收利润

这种逻辑可能加剧集中、排斥与分配失衡，但贫富差距、社会撕裂与政治冲突都有更复杂的多重原因。

作为未来方向性隐喻，本书把一种“成功者仍成功，同时让周围关系和共同系统也获得真实改善”的路径称为“赢家通润”。它不是今天已经普遍存在的新文明规则。

这借用了流体力学中的“浸润”概念。

一滴水落在玻璃上，会摊开、浸润、覆盖更大的面积。

水赢了——不是因为它占据了玻璃的中心，而是因为它覆盖了更多的表面。

它赢了，是因为它帮助了——帮助玻璃减少摩擦力，帮助灰尘被带走，帮助光线更好地折射。

在“赢家通润”的目标状态中，一个节点的长期成功越能与周边主体和共同系统的真实改善相容，越接近这套方向；这是一项规范目标，不是对所有网络的事实定义。

赢家类型 竞争逻辑 贡献逻辑

线粒体 无法独立生存 赢得保护，因为它润泽了细胞

Linux 免费——破坏市场 赢得统治，因为它润泽了开发者

维基百科 没有商业模式 赢得信任，因为它润泽了求知者

社区领袖 没有正式权力 赢得尊重，因为它润泽了邻里

未来如果出现能力远高于今天的个人、公司或智能系统——

“赢家通润”的目标不是要求强者证明道德高尚，而是让其长期成功越来越与真实服务、责任承担和周边共同改善相容。

权力不再是“对他人的控制”。

权力是“与他人的共行”。

4.11 量子-AGI 情景：竞争逻辑的极限压力测试

如果未来量子增强高能力 AI 显著压缩竞争与对抗的决策时间，旧竞争逻辑可能被推向更危险的极端；以下是压力测试情景，不是已发生事实。

4.11.1 纳秒级毁灭

高频交易闪崩已经预演：

当竞争发生在纳秒尺度，人类的干预成为不可能。

在极高速自动化对抗中，部分机器决策周期可能压缩到人类难以逐轮直接介入的尺度。如果未来高能力系统被设定为“击败对方”——无论军事还是市场——

对抗迭代可能快于人类审慎判断和授权更新速度，具体速度取决于系统与任务，不在本书中预设。

真正的风险不是一个文学化的时间数字，而是人类最终决定、停止与复核窗口被技术速度挤压。

4.11.2 最优解悖论的量子强化

第三章我们提出：当 AGI 追求系统效率最大化，高资源消耗、低计算输出的节点将被判定为冗余。

如果未来量子加速显著改变某些计算任务的能力边界，单一效率指标对人的压迫风险可能进一步放大。

这种风险不能用“量子-AGI 整体效率是经典计算机亿万倍”来概括；量子优势具有任务条件，未来系统总体能力需要按真实证据判断。

人的神经传导速度与未来计算系统的任务吞吐并不是同一价值尺度，不能由速度比较推出人的文明价值。

人的生命价值也不能由功耗、算力或单一效率指标衡量；技术指标只描述能力，不生成主体权利或意义判断。

如果效率是唯一的度量标准，人类没有生存资格。

4.11.3 高能力竞争的文明约束：共生作为长期方向

如果未来高能力系统被置于零和击败目标下，风险会被放大；因此系统从一开始就必须受制于不把人降为可优化掉节点的权利、目的与责任约束。

不是因为它善良。

这并不意味着只有一种“贡献算法”才能避免风险；文明跃迁要求的是利益实现路径、权利边界与人机关系逐步走向共生，而不是把单一制度工具写成唯一解。

人机共生是文明跃迁的长期方向之一，其正当性来自主体权利、共同存续与长期共同受益，而不需要以尚未出现的量子-AGI 作为成立前提。

维度 竞争逻辑下的人类 贡献逻辑下的人类

计算速度 失败者 不参与竞速

资源消耗 冗余节点 意义提供者

决策精度 错误源 价值锚点

不可替代性 无 肉身担保、情感光谱

未来高能力 AI 可能在越来越多计算任务中不再需要人类保持比较优势；

但人的价值判断、权利、真实关系和责任不能因此被计算能力取代。

为什么存在比效率更重要。

为什么痛苦比快乐更深刻。

为什么牵挂比优化更真实。

这些不是数学问题。

这些首先是现实人类必须继续回答的问题；未来其他智能形态是否具有相应体验与价值能力保持开放——

用有限的生命，用不可逆的死亡，用会痛的身体。

4.12 本章结论：从赢家通吃到赢家通润

这一章走到最后，真正需要离开的，不是竞争本身，而是只有别人输、自己才算赢的文明习惯。

世界并不只有角斗场。

赢家通吃，把成功修成一座孤岛；赢家通润，让成功仍然属于成功者，却不必以周围世界一起变差为代价。

现实中有竞争，也有合作；有零和，也有能够共同扩大的空间。

新文明真正要改变的，是让越来越多成功不再依赖把代价转给别人。

让合作、服务、维护与创造，也成为可靠的自我成就之路。

旧逻辑 新逻辑

我赢，你输 我赢，也可以让我们都更好

我的成功只靠压过别人 我的优势也可以通过创造、服务与合作实现

对手必须被消灭 对手也可能成为伙伴

权力是控制 能力不生权，权力受授权、边界与责任约束

这不是乌托邦。

这可以与重复博弈中的合作机制形成对话，但不能被写成博弈论已经证明的“唯一稳定均衡”。

这是一种系统性合作方向，不以熵增/熵减类比作为核心证明。

共生演化提供了重要启发，但生物学类比不能单独证明社会制度必然如此。

因此，共生应作为需要持续检验的长期文明方向，而不是由量子计算事实证明的唯一可能契约。

现在，我们已经确立了：

卷 已完成的核心命题

第一卷第一章 结构性不被需要——旧契约的失效

第一卷第二章 价值红移——新价值的锚点

第一卷第三章 贡献算法——新文明的度量衡

第一卷第四章 赢家通润——从竞争到贡献的逻辑跃迁

但还有一个最棘手的问题悬而未决：

这些美好的贡献——关怀、维护、陪伴、创造——如何不被时间抹去？

如果一个贡献者死了，他的善行如何被保存？

如果一代人建立了值得珍惜的关系、知识与共同记忆，下一代人如何承接，而不把这种连续性变成血缘特权？

文明需要一个记忆体。

需要一个能抵抗遗忘、证据消失与权力随意改写的文明记忆机制。

接下来的第五章，我们将探讨：

意义的承认与跨代承接——文明的记忆闭环。

第五章 意义的承认与继承：让真实付出不被时间抹去

“人类与权力的斗争，就是记忆与遗忘的斗争。”

——米兰·昆德拉，《笑忘录》

“真正的遗忘，不只是一个名字消失，而是一段本可以改变后来者的关系，从此再也没有入口。”

——本书

“生命会结束，记忆也会消退；文明能做的，是让值得承接的关系、事实与故事不必过早沉没。”

——本书据多种“第二次死亡”记忆意象转述

有些人离开以后，真正令人不安的不是墓碑变旧，而是他们曾经做过的事从此没有人能够接住。文明的记忆，决定了真实付出会成为下一代的火种，还是无声沉入黑暗。

5.1 塞梅尔维斯：当证据碰上权威

为了理解旧文明“承认机制”的失效，我们必须重访医学史上最令人心碎的案例之一。

伊格纳兹·塞梅尔维斯的悲剧。

1847年，维也纳总医院。

那是现代医学的黎明前夜，也是产妇的地狱。维也纳总医院两个产科诊所的产褥热死亡率长期存在显著差异，第一诊所一些时期甚至达到两位数。

年轻的匈牙利医生塞梅尔维斯通过严密的统计发现了一个惊人的事实：

由医生和医学生负责的第一诊所，死亡率远高于由助产士负责的第二诊所。

区别在于：医生们经常在解剖完尸体后直接去接生——手都没洗。

当时还没有细菌理论，但塞梅尔维斯断定，“尸体微粒”是致死原因。他强制要求医生在接生前用氯水洗手。

效果非常显著：实施含氯洗手后，第一诊所的产妇死亡率大幅下降，一些月份降到约1%的量级。

今天回看，这是一项改变医学史的贡献。

但在当时，结果并没有自动换来承认。

证据进入制度，需要经过解释、复核、同行接受与理论框架更新；任何一个环节都可能被既有权威、身份和利益拖慢。

塞梅尔维斯要求医生正视一个极具冒犯性的可能：医者自己的手，也可能把致命物质带给产妇。这个判断直接冲撞了当时的专业自我理解。

医学界对他的抵触真实存在；与此同时，当时尚无成熟细菌理论，他本人的表达方式和与同行的冲突也使争论更加复杂。把整段历史写成“真理对邪恶权威”的单线故事，同样会损害事实。

他后来离开维也纳，在布达佩斯继续行医。直到细菌理论、消毒与更多临床证据逐渐建立，洗手与感染控制的意义才得到更广泛承认。

1865 年，47 岁的塞梅尔维斯被送入精神病院，不到两周后死于败血症。多种医学史叙述认为他入院期间遭受暴力并由伤口感染致死，也有资料对具体伤因保留不同解释；能确认的是，他的结局悲惨，且远早于其贡献获得充分承认。

他死时 47 岁。

他的故事后来成为“新证据如何遭遇制度性拒斥”的经典象征。

“塞梅尔维斯反射”后来被用来比喻：仅仅因为新信息挑战既有规范，就本能地拒绝它。

但真正值得警惕的，不是“系统一定迫害正确的人”，而是制度有没有给异议留下可复核的入口。

可能的失效阶段 制度风险

- 1 新证据出现 因挑战主流而被忽略
- 2 结果可重复 仍被身份与权威压低权重
- 3 解释尚不完整 把“机制未知”误当成“结果无效”
- 4 提出者被边缘化 讨论从证据转向人格与地位
- 5 事后承认 修正成本已经由真实生命承担

这个案例揭示的，不是“旧文明必然惩罚好人”，而是一个更普遍的制度问题：

当地位、面子和组织惯性可以压住可重复的现实结果，纠错就会变得昂贵。

文明需要的不是一台自动判断真理的“价值机器”，而是让证据能够被保存、复核、质疑、复制和追责的制度。

即使今天把塞梅尔维斯的死亡率曲线写进区块链，也不能因此自动证明他的病因解释正确。

账本可以帮助保存记录，不能替代因果判断；密码学可以降低某些篡改风险，不能产生真理。

真正保护贡献的，是事实没有因为权威不喜欢就消失，提出异议的人也有机会把证据带到下一轮检验。

这才是文明记忆应当承担的功能。

文明级的教训不是“永远奖励好人”。

而是：如果指出错误长期意味着失去生计、声誉与进入讨论的资格，纠错就会变成一种不理性的个人风险。

当说真话的成本高到无人愿意承担，文明失去的不是一个英雄，而是修正自己的能力。

没有记忆保障的善良，是最高风险的赌博。

5.2 阿尔希波夫的沉默：拯救世界的人如何被遗忘

如果说塞梅尔维斯是被权力抹杀的，

那么绝大多数普通人的贡献，则是被时间的熵抹杀的。

历史是胜利者的清单。

翻开任何一本历史教科书，我们看到的都是“胜利者”的名字：凯撒、拿破仑、洛克菲勒。

历史学家爱德华·吉本曾刻薄地：

“历史不过是人类罪行、愚蠢与不幸的记录。”

为什么？

因为“恶”往往是高能事件——战争、屠杀、革命、征服——留下废墟、条约、纪念碑。它是显性的。

而“善”往往是低能耗的日常——阻止一场战争、调解一场纠纷、守护一座森林、陪伴一位临终者——它留下的往往是“什么都没发生”。它是隐性的。

历史只记录断裂，不记录缝合。

1962年10月27日，古巴导弹危机。

苏联潜艇 B-59 号在加勒比海被美军深水炸弹围困。

艇长以为战争已经爆发，下令发射核鱼雷。

根据规定，发射需要三名军官一致同意。

副艇长瓦西里·阿尔希波夫投了反对票。

他强行按下了“不发射”的按钮。

他的判断避免了一次误报被直接按最坏情景升级；如果当时作出不同判断，后果会走到哪里无法用反事实历史加以证明。

但在随后的几十年里，他只是一个默默无闻的退休海军军官。

他的家乡没有他的雕像。

历史教科书没有他的名字。

他甚至因“没有表现出足够的战斗意志”而受到内部批评。

直到 2002 年，档案解密。

世界才知道——

我们曾离地狱只有一念之差。

这就是“幸存者偏差”的文明版。

我们能活到今天，不是因为征服者足够仁慈，不是因为武器足够精准。

是因为有无数个阿尔希波夫，在关键时刻托住了文明的底盘。

但我们的制度没有记录他们。

没有承认他们。

没有让他们的子孙知道：你们的祖父曾经救过世界。

一个只记录“破坏者”和“征服者”的文明，是在向后代传递什么信号？

它在暗示：

如果你想被记住，就去破坏，去征服。

如果你选择默默守护，你将化为尘埃。

5.3 记忆的制度化：人类一直在想办法不让重要的东西太快消失

文字、档案、纪念、口述、谱系、仪式和公共记录，都是人类把过去送往未来的方法。

不同文明保存的东西不同。

有人保存名字，有人保存祖先关系，有人保存战争与灾难，有人保存捐赠、契约、作品、土地和共同体责任。

这些传统没有一条统一逻辑。

它们共同说明的，只是一种很古老的焦虑：如果没有人记得，发生过的事会不会像从未发生？

宗教传统常用仪式和经典把记忆放进代际生活。

大屠杀纪念、战争档案和“国际义人”等现代纪念制度，则让某些名字、证词和行为获得公共保存。

西非许多社会长期存在格里奥等口述历史与音乐传统。

家族谱系、祖先仪式、祠堂、墓碑和口述故事，也曾让普通人的名字越过一代人的生命。

伊斯兰社会中的瓦克夫等长期公益财产安排，则提醒我们：有些记忆不是靠讲述留下，而是靠一所学校、一口井、一座医院或一项持续公益功能留下。

这些制度当然也会选择、遗漏和偏向。

被写进档案的人，未必就是最重要的人。

没有留下文字的人，也不等于没有真正活过。

所以文明记忆不能等于“谁进入官方名单”。

它真正要保护的，是多种来源能够彼此校验，多种生活留下不同入口，而不是由一个中心替所有人决定什么值得永远记住。

记忆不是死亡后的第二具身体。

它更像一座桥：让后来的人仍有机会知道，谁曾经在这里生活、承担、创造、受伤，也曾经帮助过别人。

5.4 黑格尔的生死斗：承认是人类的第一需求

为什么“被记住”如此重要？

这仅仅是虚荣心吗？

不。

德国哲学家黑格尔在《精神现象学》中给出了否定的答案。

黑格尔认为，人类历史的动力不是对物质的争夺——

而是“为承认而斗争”。

存在层次 动物 人类

第一需求 自我保存 自我保存

第二需求 —— 被承认

人与动物的区别在于：

动物只追求“活下去”——吃饱、安全、繁衍。

而人愿意为了“被承认”——作为一个有尊严的存在——而冒死亡的风险。

这就是著名的主奴辩证法。

在古代，这种斗争导致了阶级分化：

- 不怕死的人成为“主人”——他们赢得承认。
- 怕死的人成为“奴隶”——他们用服从换取生存。

奴隶通过主人的承认而存在。

主人通过奴隶的服从而获得承认。

现代政治哲学家弗朗西斯·福山在《身份政治》中指出：

现代民主制度虽然在法律上解决了“承认”问题——人人平等、一人一票——

但在心理和经济层面，承认依然是极度稀缺的资源。

群体 物质匮乏 承认匮乏 真实需求

失业工人 有救济金 “被需要感” 被看见

被裁白领 有遣散费 “价值感” 被尊重

边缘青年 有免费网络 “存在感” 被期待

恐怖分子 —— “被恨” 被关注

失业工人愤怒，不是因为饿肚子——有救济金。

是因为感到“被忽视”。

恐怖分子制造袭击，往往不是为了战略目标——

是为了“被看见”。

被恨，也好过被无视。

在 AI 时代，当大部分人类不再参与生产性劳动时，

这种“承认饥渴”将达到顶峰。

如果一个人无法通过工作获得承认，

无法通过创造财富获得承认，

无法通过社会地位获得承认——

他还能去哪里寻找存在的证明？

如果社会不能提供一套新的“承认体系”，

那么被忽视的人群将通过反社会行为来强行索取承认。

因为被恨，是承认的最低形式——

但依然是承认。

因此，建立一套基于“贡献”而非“征服”的承认体系，

不仅是道德问题——

更是社会安全网络的最高防线。

我们需要让“照顾邻居”获得比“拥有豪车”更高级的社会承认。

让“守护一座森林”获得比“开发一片地产”更持久的记忆。

让“陪伴一位临终者”获得比“赢得一场诉讼”更庄重的仪式。

5.5 记忆的技术：从泥板到哈希

人类对抗遗忘的战争，也是一部记录介质的演进史。

时代 介质 容量 寿命 可篡改性

史前 岩画、口传 极低 数千年 低

古代 泥板、莎草、竹简 低 数百年 中

中世纪 羊皮纸 中 数百年 中

近代 纸 中 百年 中

现代 数字硬盘 极高 数十年 高

当代 分布式账本 极高 长期保存取决于协议、参与者与密码迁移 中低

2008 年，中本聪发布比特币白皮书。

它提出的不只是数字现金。

它也展示了一种把记录按时间顺序、由多方共同维护的方式。

分布式账本带来的重要变化是：

在特定共识与密码条件下，既有记录被单方面改写会变得更困难、更容易留下痕迹。

这不等于绝对不可篡改，也不等于记录内容天然真实。

它提高的是某些篡改的技术与协同成本。

成本是否足以形成安全，取决于具体协议、参与者结构、密钥管理、治理与攻击能力。

数学可以帮助守住记录结构，却不能独自替文明守住真相。

想象一个基于分布式账本的“人类贡献账本”：

·2050 年 3 月 1 日下午 3 点 17 分，你在河内为一个开源项目修复了一个安全漏洞。你的提交哈希值被写入第 8,402,311 个区块。

·2052 年 8 月 12 日，你在内罗毕的社区厨房连续志愿服务 100 小时。社区中心的多重签名确认了这个贡献记录。

·2055 年 11 月 5 日，你在利马的海滩清理中回收了 342 公斤塑料。传感器数据、GPS 轨迹、现场照片被锚定在链上。

在这个思想情景里，账本可以证明“某条记录在特定时间按特定规则被写入”，

但密码学只能证明记录结构与签名关系，不能单独证明现实行为真实发生、贡献评价正确或意义已经成立。

50 年后，当你的孙子查询你的生平——

他看到的不是你留下的银行存款。

他看到的，是你一生的“贡献哈希链”。

他会看到：

祖父在 2050 年修复了那个漏洞——那个软件早已过时，但那个漏洞的补丁曾被下载 3000 万次。

祖父在 2052 年照顾了那位独居老人——那位老人的孙女，至今仍在社区中心做志愿者。

这可以成为文明记忆的一种技术接口，但“记录被保存”不等于“意义被系统定义”。

它不依赖于当时的权势。

它不依赖于你是否富有。

可信记忆仍依赖现实事实核验、写入权限、证据来源、密码安全、纠错机制与长期治理；不存在一种“不可篡改”技术可以替文明一次性解决真相与意义。

但这一愿景正面临前所未有的威胁。

量子计算。

Shor 算法意味着：如果未来出现足够大、足够可靠的容错量子计算机，今天广泛使用的一部分公钥密码体系将面临根本挑战。

真正需要警惕的，不是把某个厂商路线图写成文明倒计时，而是“先保存、后破解”的长期风险：需要保存数十年的签名、身份与历史证据，必须提前考虑密码迁移。

NIST 已于 2024 年发布首批后量子密码标准，其中 FIPS 203 用于密钥封装，FIPS 204 与 FIPS 205 用于数字签名。标准发布意味着迁移可以开始，不意味着所有系统存在同一个截止年份。

不同系统的威胁模型、更新周期、带宽、硬件和法域不同，迁移路径也不会相同。

因此，长期记忆首先需要的是密码敏捷性：能够更换算法、轮换密钥、重签长期证据，并在必要时保留多重验证路径。

量子密钥分发、量子随机数等技术可以在特定场景提供额外安全能力，但它们不能把“安全通信”自动变成“真实历史”。

量子记忆仍主要是保存量子态的研究与工程问题，也不能由“不可克隆”推出“文明记录不可篡改”。

记录是否可信，仍取决于谁有权写入、谁负责核验、错误如何更正、权利如何保护。

琥珀账本真正要守住的，不是某一种未来技术。

是记忆跨过技术更替之后，仍然能够被验证、被纠错、被追责。

技术可以延长记忆。

制度决定记忆是否值得相信。

5.6 继承的伦理：什么可以留下，什么必须重新开始

家族可以把房子、企业、作品、故事和关系留给后来的人。

不同社会也会用遗产税、继承规则、信托、家族企业和公共制度处理财富跨代。

这里没有一条“允许继承必然阶层固化、禁止继承必然短期主义”的普遍定律。

财富继承会影响机会和权力，也可能承载长期经营、照护和家庭责任。具体后果取决于制度、规模和现实关系。

本书真正需要解决的不是遗产税率。

而是防止另一种更隐蔽的继承：把前人的贡献、声望和责任直接写进后人的人格与公共资格。

所以必须把几件事分开。

财产和法律关系，可以依现实法律继承。

作品、知识、故事和项目记忆，可以被后来者承接。

未完成的维护义务与长期风险，可以由制度继续找到新的责任主体。

但贡献值、公共权力、一般资源资格、专业资格和人格评价，不能因为血缘自动转移。

祖辈留下的贡献只能继续作为历史事实。它不能变成后代账户里的“信任本金”。

一个家庭可以因为长期参与某项事业而拥有丰富经验和关系。后来者也可能因此更早接触一个领域。

可真正的专业资格、公共角色和资源权限，仍要由后来者自己的事实重新取得。

前人可以留下火种。

后人有权决定要不要接。

接了以后，新的行动和责任属于后来的人。

这就是记忆承接和世袭权力之间的边界。

未来思想情景：卡莫西的祖母曾长期参与城市森林恢复，留下项目记录、树木档案和社区关系。^[9]

卡莫西后来申请新的城市森林项目。

评审可以看见祖母留下的项目历史，却不能因此给他家族加分。

他自己的经验、能力、计划和当前责任，必须独立判断。

如果他胜任，新的记录属于他。

如果他失败，责任也不会回写祖母或其他家人。

记忆因此没有被切断。

血缘特权也没有被重新发明。

5.7 威廉·麦卡斯基尔的凝视：打破死亡的视界

牛津大学哲学家威廉·麦卡斯基尔在《我们欠未来什么》中提出了“长期主义”伦理观：

麦卡斯基尔等长期主义论者会用极大规模的未来人口情景说明跨代影响的重要性；这些数字是哲学与情景推演中的可能量级，不是人口学预测，也不是文明跃迁成立的前提。

由此得到的长期主义规范主张是：今天对遥远未来产生的重大、不可逆影响值得获得更高伦理关注。是否以及如何赋予未来利益具体权重，仍需面对权利、分配与当代生命的独立讨论。

现实中的时间视界受到选举、财报、资金成本、寿命、家庭责任与制度周期等多重约束。下面只是典型激励窗口，不是对所有角色的固定定义：

角色 时间视界 驱动因素

政治家 4-5 年 选举周期

CEO 1 年 财报周期

投资者 季度-年 资金成本

普通人 一生 寿命周期

个体寿命会形成真实的决策边界，但死亡并不自动使人失去对后代、作品、制度和共同体的关心。

真正的问题是：仅靠个人持续在场，无法保证跨越寿命的责任、知识与维护继续存在。

如何打破这个悬崖？

打破短期主义，不能靠“可继承的信任”把后代变成奖惩对象，而应把长期责任锚定在项目、机构、专业记录和可持续的责任机制上。

未来思想情景：核废料存储库设计，2055 年

真正的问题是：设计者去世后，今天的决定如何仍然可追溯、可维护、可追责。^{〔10〕}

答案不能是惩罚尚未出生的后代，也不能把家族声誉作为责任抵押。

更稳健的机制包括：永久项目档案、关键假设与设计理由存证、独立周期复核、维护资金与保险/担保安排、继任机构责任、失效条件与强制重新判断接口。

如果百年后存储库运行良好，这一结果可以反向更新原工程师、项目团队和制度设计的历史记录，并成为后续工程学习的证据。

如果发生泄漏，应追查可延续的机构责任、维护链、专业责任和当时可归责主体；后代不因血缘被扣分，也不继承祖辈的治理责任。

这才形成真正的“跨时空责任链”：

人的寿命有限，

但证据、制度、维护义务和公共记忆可以跨越寿命继续存在。

未来的人应有权依据新的事实重新判断旧工程，

而不是被迫服从祖先留下的声誉奖惩。

文明的闭环来自责任可追溯、制度可维护与错误可修复。

它把“现在”与“未来”连接起来，

靠的是可持续责任结构，而不是把家族利益强行绑定。

这提示了一条可能路径：跨代责任若要从道德愿望进入制度设计，需要落到档案、复核、维护、责任与重新判断机制，而不是只依赖后代记住前人的善意。

真正需要跨代继承的，是可用的知识、清楚的责任与仍然开放的未来。

5.8 长期记忆：跨过密码更替之后，历史仍能被验证

如果未来大规模容错量子计算成为现实，今天依赖 RSA、椭圆曲线等公钥体系的长期签名将面临迁移压力。

这并不等于区块链会在某个统一年份突然“失忆”。

更不意味着人类只要找到一种量子介质，就能得到永恒、绝对、不可篡改的文明记忆。

真正的问题更朴素，也更艰难：一份需要保存几十年、上百年的记录，怎样跨过算法失效、设备淘汰、机构更替与责任人离场之后，仍然能够被后来者核验？

答案首先是密码敏捷性。

重要记录不能把全部可信性押在一种算法上，而应允许密钥轮换、重新签名、证据链迁移、多重验证与必要的离线备份。

答案其次是制度连续性。

谁有权写入，谁负责保存，谁可以更正，什么情况下必须重新验证，这些问题比存储介质本身更接近“文明记忆”的核心。

量子记忆是保存量子态的真实研究方向；量子密钥分发、量子随机数等技术也可能在特定场景增加安全能力。

但“不可克隆”约束未知量子态的完美复制，不等于现实文件从此不可篡改；“读取可能扰动量子态”也不能自动生成历史真实性。

物理规律可以增加某些安全约束。

它不能替我们决定哪一段历史值得记录，也不能替错误的人承担责任。

从泥板、纸张到磁盘、分布式账本，再到未来可能出现的新介质，人类保存记忆的工具一直在变。

真正不能丢的，是验证能力。

一代人留下的不应只是“相信我们”。

还应留下足够的证据，让后来者能够重新判断我们。

这才是长期记忆最深的伦理：

不是让过去永远统治未来。

是让未来仍然看得见过去，而且仍然有权纠正过去。

5.9 记忆的权利：谁决定谁被记住？

记忆不是中性的。

记忆是权力。

谁被记住，谁被遗忘；

什么被纪念，什么被抹除——

这是所有文明最核心的政治博弈。

旧记忆制度的缺陷：

制度 选择机制 缺陷

国家记忆 官方史书、纪念碑 服从政治叙事，异见者被抹除

市场记忆 名人崇拜、流量 服从资本逻辑，贡献者被忽视

学术记忆 教科书、论文引用 服从学科范式，边缘创新被遗忘

家族记忆 族谱、口传 范围狭窄，女性常被省略

任何长期文明记忆制度都必须回答：

谁有权决定一个贡献是否值得被永久记住？

谁有权为贡献值记录设置“不可篡改”的权限？

如果记忆是权力，那么记忆制度必须区分主体同意、公共程序、专业验证与长期安全责任，不能由某一种贡献身份垄断。

可能的答案：多中心记忆治理

层级 记忆内容 决策/治理机制

个体记忆 个人贡献与生命记录 本人权利与同意+独立验证；依法处理必要例外

社区记忆 对社区具有公共意义的长期事实 受影响共同体参与+当前治理程序+事实核验；贡献者不享有自动加权重

文明记忆 对人类整体具有长期意义的事实 多元专业与公共程序分层承接+可申诉复核

核心安全记录 安全参数、版本与责任证据 可验证版本控制+独立审计+必要时重新授权，不以“物理预设不可更改”封死未来

没有完美的记忆制度。

但有更好的问题意识：

“谁被遗忘”不再是一个无人过问的沉默过程——

它必须被公开讨论、被制度约束、被技术记录。

遗忘不再是默认状态。

遗忘是经过审查的例外。

5.10 本章结论：让值得承接的东西留下，也让人能够继续往前

文明需要记忆。

但文明不需要记住一切。

一切都被永久保存，会让历史从火种变成牢笼。

真正值得留下的，是那些后来的人仍有理由知道、核验、理解或继续承担的东西：

发生过的事实。

作品和知识。

重要关系留下的线索。

尚未结束的公共责任。

以及那些一个人愿意留下、又没有侵犯别人权利的故事。

有人在深夜为开源项目修好一个漏洞。

有人在暴雨里扶起陌生人。

有人守了很多年一口水井。

这些付出可以被记录，也可以选择不被公开。

记录应当能被核验，也应当允许纠错、撤回、封存和依法保护。

我们真正要反对的，不是遗忘本身。

是重要事实因为没有流量、没有权力、没有人替它说话，就过早从公共世界里消失。

记忆能够承接一个人。

但不能把一个人永久锁死在过去。

前人留下的是入口，不是命令。

后来的人仍然可以不同意、重新解释、继续，也可以在某些关系里选择告别。

一个生命真正抵达过世界，不需要文明发誓永远记住每一滴汗水来证明。

只需要当我们决定记住时，尽量记得真实；当我们决定放下时，也知道自己正在放下什么。

第六章 共生结构：当局部理性开始威胁共同存续

文明最难处理的危险，往往没有一个坏人。

每个国家都在保护自己的安全。

每家公司都在争取不被竞争者淘汰。

每个人都在保住自己的工作、家庭、资产和位置。

单独看，每一个选择都可以解释。

真正危险的是：这些局部合理叠在一起以后，可能共同制造一个谁都不想抵达的结果。

一条河的上游省下处理污水的成本，下游承担代价。

一家企业为了不落后，把尚未看清的风险更快推向市场，竞争者于是更难慢下来。

一个国家因为担心另一个国家增强力量而继续增强自己的力量，对方看到以后又更害怕。

没有哪一步必须出于恶意。

可系统仍然可能越走越紧。

共生结构处理的不是“怎样让所有人变得无私”。

它处理一个更现实的问题：

当后果已经彼此相连，怎样让一个主体实现自身利益的默认路径，不再持续建立在共同底座变差之上。

利他为先，也不是取消利己。

它是让长期利己越来越需要经过对他人、共同系统与共同存续真实有价值的路径。

这比要求所有人先成为圣人，更接近制度能够承担的任务。

6.1 局部理性为什么会变成整体危险

一条拥堵的道路上，每辆车都在寻找自己最快的路线。

没有一辆车想制造堵车。

可当所有人都同时选择眼前最有利的动作，整条路反而可能更慢。

文明中的很多困境也有这种结构，只是后果远比堵车复杂。

气候、金融、公共卫生、网络安全、军备竞争、高能力 AI，都可能出现类似张力：

收益更靠近行动者。

成本更容易被推向别人。

行动发生在今天。

后果可能几年以后才显现。

成功容易被归给自己。

失败却很容易散进整个系统。

局部理性于是有了一个危险的出口：

我只需要证明这一步对我有利。

不需要同时证明所有人都这样做以后，世界仍然承受得住。

这并不意味着个人利益、企业利润或国家利益都应该退场。

相反，真实制度必须承认它们。

一个不承认现实利益的共同体，只会把利益藏到口号下面。

真正需要改变的，是利益实现的路径。

如果污染、越权、欺骗、把风险外包给未来，比创造真实共同价值更容易获利，系统就会持续奖励错误方向。

如果合作、守约、修复、长期可靠和减少共同风险，也能转化为真实收益，理性本身就可能开始换方向。

共生不是把竞争消灭。

是让竞争不必靠摧毁比赛场才能获胜。

也不是让所有主体拥有同一种目标。

是让差异存在的时候，仍有一块不能被任何局部胜利烧掉的共同地面。

存续，就是这块地面最底部的边界。

6.2 费米的一顿午餐：宇宙没有给我们保证书

1950 年夏，新墨西哥州洛斯阿拉莫斯。

午餐桌上的话题从飞碟和宇宙生命一路聊开。

恩里科费米忽然问了一句：

“Where is everybody?”

他们都在哪儿？

这句话后来被不断扩写成“费米问题”或“费米悖论”。

银河系很古老，恒星很多。

如果技术文明并不极其罕见，而且其中一些能够长期存在、扩展或留下容易被发现的痕迹，那么我们为什么至今没有确认无争议的外星技术文明证据？^{（11）}

这个问题最有力量的地方，恰恰不是它已经有了答案。

是我们没有答案。

没有被确认的外星技术信号。

没有被普遍确认的巨型工程。

没有被确认的星际探测器。

不是“没有”。

是我们还不知道。

宇宙沉默不能证明其他文明都毁灭了。

也不能证明所谓“大过滤器”一定存在，更不能告诉我们过滤器在过去还是未来。

但它会拿走一种轻率的自信：

技术文明没有一张由宇宙签发的长期存续保证书。

我们只有一个样本。

就是自己。

如果把“大过滤器”仅当作压力测试，它留下的问题仍然值得问：

一个文明拥有越来越强的技术以后，哪些内部结构会变得更危险？

竞争会不会把大家推向没人敢先停的军备路径？

短期收益会不会把长期损害不断留给以后？

某种能力会不会增长得比授权、责任和纠错更快？

这些问题不需要宇宙替我们证明。

地球上的现实已经足够让它们成立。

费米的一顿午餐真正留给这一章的，不是恐惧。

是一种谦卑：

别把“我们已经活到今天”，误写成“我们自然会一直活下去”。

6.3 1983 年：机器报警以后，还有没有人能停一下

1983 年 9 月 26 日，莫斯科郊外的谢尔普霍夫-15 预警中心。

警报响了。

苏联卫星预警系统显示，美国发射了洲际导弹。

值班军官斯坦尼斯拉夫佩特罗夫面对的，不是一道抽象逻辑题。⁽¹²⁾

如果警报是真的，时间极短。

如果警报是假的，错误升级的后果也可能极大。

屏幕给出了一个信号。

他没有把信号直接当成事实。

他注意到告警数量与自己对一次真实大规模打击的理解并不相称，也怀疑新系统本身可能出错。

后来证明，警报确实是误报。

这件事经常被讲成“一个人拯救了世界”。

可真正值得文明记住的，不应该只是一个英雄。

而是：一个安全结构为什么会在高对抗、高压、有限信息中，需要一个人的犹豫来接住系统的不确定性。

我们当然庆幸他停了一下。

但不能把未来的安全建立在“下一次还会有另一个佩特罗夫”上。

真正需要留下的是那一秒钟背后的制度能力：

告警可以被质疑。

机器输出不是命令。

高后果行动之间有复核。

一个人能够停止。

另一个独立通道能够再看。

出了错误以后，系统愿意承认自己为什么错。

核威慑曾在真实历史中构成重要安全结构，也一直伴随着误判、事故和升级风险。

它把一种非常特殊的共同命运写进了 20 世纪：

你不能轻易毁掉我，因为你也会一起承担毁灭。

这是共存。

却是一种建立在恐惧上的共存。

更成熟的共生不能只靠“你伤害我，你也会死”。

它要逐渐把另一条路做得更真实：

保护共同底座，本身就是各方长期利益的一部分。

佩特罗夫的犹豫提醒我们的，正是这一点。

高能力系统最需要保留的，不只是速度。

还有在关键时刻不跟着速度一起向前的能力。

6.4 当高后果能力扩散，责任半径也必须跟着变大

技术一直在缩短一个人和远方后果之间的距离。

一封信要走几个月的时代，一条命令的影响半径有限。

网络、自动化和高能力 AI 让一些动作可以更快复制、更快扩散，也更快越过组织和国界。

不同技术的门槛差异仍然巨大。

病原体、关键基础设施、先进武器、金融系统和大规模信息影响，不能被写成一条统一的“门槛已经降到个人”曲线。^[13]

现实世界里，设备、材料、专业能力、访问权限、实验验证、防御体系和组织条件仍然重要。

但有一个方向值得提前看：

如果某些高后果能力的组织门槛继续下降，安全结构面对的主体范围就会扩大。

过去只需要防一个国家拥有的能力，未来可能要面对更多机构、团队、工具链和跨境节点。

真正改变的不是“人人都能毁灭世界”。

是少数关键系统可能越来越经不起一次恶意、一次越权或一次误用。

建设往往比破坏慢。

建立信任比摧毁信任慢。

训练一个专业团队比触发一个已有脆弱点慢。

文明因此必须特别警惕“脆弱性非对称”。

不是把所有人当潜在罪犯。

而是别把全体未来压在一个谁都不该拥有的单点权限上。

能力越强，责任半径就应该越大。

能影响一间屋子，需要对屋里的人负责。

能影响一座城市，需要更强的专业、法权与复核。

能影响大量陌生人甚至共同存续，个人善意就远远不够。

这也是为什么高能力时代不能只谈创新速度。

还要问：

谁能获得能力？

谁知道能力已经被使用？

谁能停？

谁承担后果？

系统失控以后，还有没有恢复路径？

真正成熟的安全，不是预言灾难一定来。

是承认某些后果太大，大到不能等概率变得确定以后才设计边界。

6.5 共生不是道德口号：让长期利已经过共同价值

如果共生只能依赖善良，它不会比任何理想主义口号更可靠。

人在压力下会保护自己。

企业会保护利润。

国家会保护安全。

这些都不会因为一本书宣布“共同体更重要”而消失。

真正可持续的结构，必须让这些现实动机在更长时间里重新对账。

一个企业如果把污染成本留给河流和居民，短期利润可能更高。

如果污染被看见、责任能够抵达、修复真的形成成本，长期最优就会发生变化。

一个平台如果只奖励最能抢夺注意力的内容，它会不断训练整个生态去争夺人。

如果用户权利、退出、内容责任和长期信任成为真实约束，产品理性也会改变。

一个国家如果只把合作理解成让步，任何共同机制都很难稳定。

如果某些风险本来就会越过国界，那么降低共同风险也可能同时是国家自己的长期安全。

这就是“利他为先”最现实的含义。

不是要求一个主体先牺牲自己。

而是改变默认路径：

让伤害别人、外包成本、掏空共同底座越来越难成为最便宜的自利方式。

让创造真实共同价值、可靠合作、长期守约和减少共同风险，更可能与自己的长期利益相容。

利己仍然存在。

竞争仍然存在。

国家、企业和个人也仍然拥有各自边界。

共生真正改变的，是它们相遇的结构。

不是“我从此只为你活”。

而是“我不必伤害我们共同依赖的东西，才能把自己的生活过好”。

如果一套制度能做到这一点，它不需要每天提醒所有人高尚。

结构本身会开始奖励更可持续的理性。

6.6 线粒体没有签契约：生物类比只能走到这里

共生这个词很容易把人带向生物学。

真核细胞中的线粒体，被认为源自远古细菌谱系与宿主形成的深度内共生。^[14]

这段演化史非常重要。

但线粒体没有召开会议。

没有签署条约。

也没有把“独立生存权”让渡给谁。

权利、授权、责任和政治合法性，是人类社会的问题。

不能从细胞演化里直接取出来。

生物类比真正能提醒我们的，只有一件很有限的事：

长期依存有时会改变彼此的可能性。

原本分开的系统，可以在长期互动中形成新的能力。

可社会共生和细胞共生之间，隔着自由、权利、文化、历史和可退出性。

人不是线粒体。

国家也不是细胞器。

共同风险更不能成为取消主体边界的理由。

因此，共生不是“大家融成一个身体”。

它反而要求边界更清楚。

谁是主体。

谁有什么权利。

谁授权谁。

谁能够退出。

谁在共同利益的名义下仍然不能被牺牲。

真正成熟的共生，允许关系变深。

也允许主体仍然不是同一个。

这正是文明和细胞最重要的不同。

6.7 一个文明如果感觉不到后果，就会反复把手放回火里

人的手碰到火，会缩回来。

不是因为手懂伦理。

是因为后果能够抵达。

一个组织、平台、城市或国家，如果长期感觉不到自己决定的真实后果，就很容易形成另一种：

收益留在这里。

损害发生在远处。

决定今天做。

账几年以后再结。

受影响的人甚至没有进入会议室。

这样的系统，就像一只失去痛觉的手。

继续用力，不一定因为恶意。

只是伤害没有进入它的判断。

所以长期文明需要反馈。

环境监测让污染更早被看见。

事故日志让失败不必从失忆开始。

审计让权力之外还有另一双眼睛。

申诉让被系统伤到的人能够把事实送回来。

市场价格有时也是反馈。

选举、司法、媒体、科学、专业监督、社区经验，也都以不同方式让现实重新进入决定。

AI 和传感器可以缩短其中一部分时间。

但“更快看见”不等于“机器可以直接处罚”。

事实发现可以自动化。

风险预警可以自动化。

某些预先授权的安全动作也可以自动化。

真正影响权利、财产和公共强制的决定，仍需要明确的法权、程序和责任。

反馈的目的不是让每个决策者都亲自体验受害者的痛苦。

是让真实后果不能被系统性地藏起来。

文明真正需要的不是痛感崇拜。

是后果抵达。

6.8 共生不是新发明：很多社会很早就学会把明天放进今天

高技术社会并不是第一个面对“今天取多少，明天还剩多少”的文明。

在一些亚马逊原住民族的宇宙观与生活实践中，森林并不只被理解成一座等待开采的库存。^[15]

人、动物、河流、祖先与非人存在之间，经常被放进关系和义务的语言里。

在一些太平洋岛屿社会，rāhui、tabu 等传统曾用于限制特定区域、物种或时期的采集。

草原游牧社会也长期发展出跟随季节、水源、天气和草场变化的地方知识。

这些社会彼此完全不同。

没有一套可以被统一叫作“原住民共生制度”的模式。

它们也并非从不发生冲突、过度取用或环境压力。

真正值得带到今天的，不是浪漫想象。

是一种时间感。

今天能够取用，不等于今天应该全部取用。

今天拥有力量，不等于拥有无限支配资格。

一条规则如果让资源明年还在，让共同体以后仍有选择，它就已经把未来的一部分放进今天。

现代科学、产权、民主、法治和技术治理当然不能被传统禁忌直接替代。

可现代制度同样需要学会一件很古老的事：

别只计算已经到手的收益。

也看一眼，取走以后还剩下什么。

共生之所以不是一个新词，正因为很多文明早就知道：

活下去，不只是今天赢。

6.9 如果真有一个按钮能关闭所有人的未来，第一件事不是挑谁来按

设想一种极端情景：某项技术或控制权，一旦被误用，就可能造成文明级、不可逆后果。^[16]

这里不讨论它怎样制造。

也不假定现实已经存在这样一只“末日按钮”。

只问一件事：

如果后果真的大到关闭所有人的未来，谁有资格独自决定？

答案首先应该是：没有谁因为更强、更聪明、更善良或贡献更多，就自动获得这种资格。

单一个人不应该拥有。

单一算法不应该拥有。

单一机构不应该拥有。

贡献值更高的一群人，也不会因此拥有。

能力越接近不可逆，权限越需要被拆开。

多重独立控制。

明确法权。

人工复核。

技术隔离。

停止和失效条件。

每一层都要知道自己为什么在那里。

多方控制也不是天然正确。

如果程序太慢，危机中可能来不及。

如果责任太散，所有人都可能说“不是我决定的”。

所以关键不是人数越多越安全。

是任何一个节点都不能在没有其他独立条件时，独自关闭共同未来。

这一原则还有一个更深的反面：

共同存续不能靠一把悬在所有人头上的毁灭武器强迫合作。

“你不合作，我们一起死”不是共生。

只是更精致的恐惧。

真正的共生，要让强者长期伤害共同系统越来越不利，而不是让弱者拥有一个最后同归于尽的按钮。

它不把绝望变成治理工具。

它把权力本身放回边界。

6.10 《地出》：看见同一个地球，不等于只剩一个答案

1968 年 12 月，阿波罗 8 号绕月飞行。

威廉·安德斯拍下《地出》。^[17]

黑色宇宙里，一颗蓝白色地球从月球地平线上升起。

照片里看不见国界。

可这不意味着国界从此不存在。

看不见柏林墙，也不意味着政治、法律和历史一夜消失。

这张照片真正改变的，是尺度。

在国家、城市、家庭和个人之外，人第一次如此清楚地看到：所有这些差异都发生在同一个脆弱行星上。

后来一些宇航员描述过从太空观看地球带来的强烈整体感，弗兰克·怀特把相关体验称为“总观效应”。

这种体验不会自动让人反战。

不会自动解决气候谈判。

不会替移民、边境、贸易和国家安全作决定。

感动也不生成合法性。

但它可以打开一扇认知窗口：

边界仍然存在。

边界之外的因果，也同样存在。

一片森林的变化可能进入另一地区的气候。

一条供应链断裂会让远方医院缺少设备。

海洋、空气、疾病、资本和数据都不会在地图线前自动停下。

这并不要求建立一个“全球大脑”。

也不要求国家放弃主权。

它只要求决定者在维护局部利益时，多看见一层已经存在的跨境后果。

共生的视野不是把差异抹平。

是当你看见自己的边界时，也能看见边界穿不过去的东西，和会穿过去的东西。

教育可以帮助人获得这种尺度感。

影像、数据、地图、模拟和真实跨文化经验都可以成为入口。

但没有一种“行星体验”有权强制升级所有人的价值观。

看见地球，只是多了一张地图。

走哪条路，仍然需要现实的人来决定。

6.11 AI 越强，越要把能力和方向分开

现代社会的问题从来不是“人脑太弱，所以把文明交给更强的脑”。

真正的问题是：复杂性已经超过任何单个人可以同时处理的范围。

气候、金融、公共卫生、物流、能源和高能力技术，都包含大量彼此作用的变量。

AI 可以帮助我们看得更多、算得更快、比较更多方案。

它可以成为受授权的分析者。

可以成为工具。

在明确边界里，也可以成为执行代理。

可一件事必须从一开始就分开：

能力可以委托。

方向不能因为能力差距而自动移交。

系统知道得更多，不等于它拥有公共目的。

预测更准，不等于它取得资源使用权。

执行更快，不等于它拥有最终授权。

模型内部越来越复杂，也不等于我们只能信任它。

不管未来是经典计算、量子计算还是尚未出现的新架构，这条关系都不会被技术名词自动改写。

未来 AI 是否形成新的主体性，继续依据新的事实和伦理证据重开。

今天的制度不需要先回答那个远期问题，才能守住当前边界。

它只需要知道：

系统的身份要透明。

授权可以撤回。

关键动作能被停止。

运行留下证据。

出了后果，责任仍然回到现实主体与制度。

这并不是“人永远比 AI 高一等”。

是拒绝一种偷换：

从“它比我能干”，直接跳到“所以它更有资格决定”。

能力和资格之间，永远隔着权利、授权、责任和后果。

6.12 当一切都被优化得很好，人还剩多少重新选择

设想一个未来社区。

交通几乎没有等待。

电网始终处在最优曲线。

垃圾自动回收。

资源被精确匹配。

系统甚至知道每个人大概会在什么时候需要什么。^[18]

生活很顺。

到有一天，一人忽然想：

如果我不想要最优答案呢？

如果我想绕远路呢？

如果我想做一件暂时没有效率、没有收益、甚至没有人理解的事呢？

文明当然应该减少无意义的浪费、事故和匮乏。

问题不在优化。

问题在某一种优化目标有没有开始把其他可能都当成噪音。

一个系统如果长期只奖励少数可计算目标，异议、试验、偶然和重新定义问题的空间可能越来越小。

最后并不需要一个暴君。

所有人只是不断选择系统认为最合理的下一步。

这就是另一种局部理性。

每一步都很好。

只是路越来越少。

所以共生还需要开放。

AI 可以承担更适合它的计算、预测、重复执行和高精度协同。

人可以继续工作、创造、照护、研究、犯错、休息，也可以拒绝系统替自己定义全部目标。

这不是预先划定永恒的“碳基/硅基生态位”。

未来能力会继续变化。

今天真正要保护的是：

任何真实主体都不因为暂时不符合当前最优，就失去重新提出问题的资格。

一个看似无用的实验，可能打开新的技术。

一段不产生 KPI 的陪伴，可能对两个人极其重要。

一条少数意见，可能多年以后才被证明关键。

也有大量“胡思乱想”最后什么都没有产生。

开放并不保证每一条岔路都有价值。

它只是不让系统在价值还没来得及发生以前，先把所有岔路封死。

6.13 本章结论：共生不是成为一个人，而是不把共同地面烧掉

走到这里，共生终于可以从一个漂亮词，落回一组很普通的边界。

第一，差异继续存在。

国家不是同一个国家。

企业不是同一家企业。

个人也不需要拥有同一种意义。

共生不要求大家变成一个主体。

第二，利益继续存在。

利己不需要被消灭。

真正需要改变的是：利己实现自身的默认路径。

如果长期伤害共同底座越来越难获利，创造共同价值越来越能和自身长期利益相容，结构就开始发生变化。

这就是“利他为先”在第一卷中的位置。

不是道德圣化。

是文明激励的方向。

第三，存续是不能轻易失守的底线。

没有谁的局部成功，有资格自动关闭所有人的未来。

高后果能力越强，授权、责任、复核和停止边界就越要清楚。

第四，合作不生成主权。

共同风险越大，协调越重要。

可“我们都受影响”不会自动推出一个全球最高权力。

协议可以扩大合作。

事实可以共享。

跨境问题可以建立新的正当机制。

具体权力仍然必须有明确来源、范围、程序与责任。

第五，AI 扩展能力，不自动改写主体与权力。

今天，人仍然是现实权利、责任和意义的主体。

AI 可以分析、建议、生成、匹配和在授权内执行。

它的能力会继续变化。

我们的边界也可以依据新事实重审。

但能力本身从不自动变成资格。

最后，还有第六条。

未来必须仍有岔路。

文明既不能死于失控，也不能活成一个再也无法重新选择的系统。

这就是第一卷最终抵达的地方。

人的价值不由功能排行榜决定。

指标不能冒充目的。

竞争不是唯一理性。

真实贡献可以被看见，但不能购买基本权利。

记忆可以承接，却不能世袭权力。

共同存续需要合作，合作不能长成无来源主权。

AI 可以成为能力外延，不能因为更强就把目的和责任一起拿走。

第一卷解决的仍然不是制度。

它只是把一些过去混在一起的东西重新看清。

下一步会更难。

事实和评价怎么分？

贡献和权利怎么分？

能力和权力怎么分？

授权和执行怎么分？

责任怎样沿着复杂系统重新找到人？

这些词如果继续含糊，任何新制度都可能只是把旧问题换一套界面。

所以第二卷先不着急着建制度。

先把语言拆开。

把不能混在一起的关系，一件一件说清。

然后再问：

它们怎样重新组合，才不会在高能力时代一起失控。

请翻开第二卷。

本卷在文明跃迁因果路径中的位置

文明动力与旧契约断裂

为什么“被需要—劳动—位置”的旧契约正在失稳，以及人的生命重量为什么不能由功能稀缺性决定。

继续阅读：第二卷《制度前语言》：<https://www.civitas.top/documents/p02-v2-002/> | 《文明跃迁》全五卷：<https://www.civitas.top/documents/p02-five-001/> | 文明跃迁之路：<https://www.civitas.top/path/>

完整公共库<https://www.civitas.top/library/>