

· 人工智能与未来社会(二十五) ·

迈向“负责任”的知识

——AI 时代知识生产流动的失序风险及其“敏捷—包容”治理

安小米

【内容摘要】 生成式人工智能的崛起推动人类知识生产流动呈现出“人机协智”的新范式，知识的生产与流动也在生产主体、生产模式、形态载体、流动方式、循环速度五个方面呈现出新的特性。AI 为知识生产和流动赋能的同时，也在真实性、公平性、主体性、权力性和文明性等层面带来了失序风险。面对 AI 导致的知识生态复杂性与既有单一治理模式的失灵，需探索一种“敏捷—包容”协同的混合治理范式。“敏捷—包容”视角下的“知识四用”治理框架，通过“知识可用—知识有用—知识易用—知识善用”的梯次递进、相互支撑，将“负责任”从抽象理念转化为可操作、可落地的治理体系。该框架旨在通过各层级要素的协同耦合，系统回应 AI 时代的知识失序风险，为构建人机共生背景下的健康知识生态提供清晰的行动蓝图。

【关键词】 知识生产流动 “负责任”的知识 “敏捷—包容”治理 “知识四用”框架

【作者】 安小米，中国人民大学智慧城市研究中心主任、吴玉章讲席教授、数据工程与知识工程教育部重点实验室(中国人民大学)大数据治理与知识管理团队负责人。(北京 100872)

【基金项目】 国家社会科学基金重大项目“我国政府数据治理与利用能力研究”(20&ZD161)

问题的提出：AI 时代的知识何以“负责任”

从 ChatGPT、DeepSeek 等生成式人工智能的“现象级”应用开始，人类知识的生产与流动正以前所未有的方式迈入“人机共生”的新纪元。在此过程中，知识这一贯穿人类文明演进的核心要素不再仅仅是由人类心智创造，经由语言、文字、印刷、电子和互联网逐级传递的静态产物，而是演变为由人机共生、动态交互、实时演化的复杂系统。在这一系统中，AI 不仅是知识传播的加速器，更是知识生产的协作者、解读者甚至创作者，推动知识生产从“人类独白”走向“人机协智”。





然而,效率的提升与流动的加速也带来前所未有的系统性风险。知识生产流动的“失序”已不再是传统意义上的信息过载或真假混杂,而是演变为一场对真实性、公平性、主体性、权力结构与文明延续构成全面挑战的认知危机。在此背景下,我们无法再沿用传统的治理逻辑,无论是单一的政府监管,还是纯粹的技术自治,抑或是片面追求效率或公平的单一路径,都已难以应对 AI 所带来的知识生产流动的复杂性与不确定性。因此,本文旨在探讨,在 AI 时代如何构建一个有效的治理框架,以确保其流动是“负责任”的,从而既能最大化其创造福祉的潜力,又能最小化其对个体、社会和文明带来的系统性风险。换言之,本文的核心关切在于,在效率与失控并存的 AI 时代,需在“敏捷—包容”协同的混合治理视角下,构建“知识可用—知识有用—知识易用—知识善用”的知识生产流动“四用”框架,以此驾驭知识洪流,使其从“无序”走向“负责任”。

从口语传播到“人机协智”:知识生产流动演进史

① 韩震:《知识形态演进的历史逻辑》,《中国社会科学》2021 年第 6 期。

② 查尔斯·范·多伦:《知识的历史》,张圆译,重庆:重庆大学出版社,2023 年,第 3 页。

③ 张兵:《关系、网络与知识流动》,北京:中国社会科学出版社,2014 年,第 3 页。

“人类因内在地具有知识内涵的劳动而与动物区别开来。”^①因此,从某种意义上说,“人类的历史就是人类知识进步和发展的历史”,^②其中记录着知识形态本身在演进过程中的重大突破与深刻变迁。从口语时代到 AI 时代,知识生产流动呈现出从依附于人,到独立于物,再到赋能于机的演进历程。

(一) 知识生产流动演进史:从口语传播到人机协智

所谓知识生产流动,是指知识从被创造、编码、传播,到被接收、吸收、应用并最终催生新知识的动态迭代过程。它包含了知识“从无到有”与“从此到彼”两个核心维度,可以被拆解为两个相互关联的部分:一是知识生产,指的是新知识、新见解、新思想的创造与系统化过程,是知识的源头和增殖环节,回答的是知识“从无到有”的问题;二是知识流动,指的是知识在不同主体(个人、组织、社会)之间,通过特定渠道和媒介进行传播、共享、吸收与应用的过程,是知识的连接和效能环节,解决了“从此到彼”的问题,确保了知识不会孤立和消亡。知识生产与流动是一个相互依存、相互促进的闭环,二者均会带来知识创新。一方面,生产驱动流动,新知识的产生是流动的起点和内容。另一方面,流动反哺生产,知识的流动促进了思想的碰撞与融合,是新知识生产的催化剂,流动过程中产生的反馈、质疑和应用场景,会直接激发和引导新一轮的知识生产。概言之,知识流动是知识生产的必要条件。^③更为重要的是,从知识运动的本质来看,知识生产本身也是一种特殊的知识流动。具体而言,知识生产是对既有知识要素进行重组、整合、演绎或超越的过程。这一过程中的知识在不同认知主体之间持续发生着输入、加工、转化与输出的“流动”。因此,知识生产的核心机制正是将既有知识通过认知活动“流动”为新知识的过程。换言之,知识生产不是流动的起点之外的一个独立环节,而是知识在更高层次上、在认知系统内部完成的一次深度流动与转化。在传统社会,知识生产与流动往往呈现出相对清晰的阶段划分——生产多发生于特定的学术或专业机构,流动则是其后向社会的扩散。然而,在 AI 时代,这种边界正变得日益模糊,AI 不仅可以自主或辅助生产新知识,同时也重塑着知识流动的渠道、速度与形态。知识的生产与流动日益融为一体,形成一个“在流动中生产,在生产中流动”的动态循环和迭代改进。因此,本文以“知识生产流动”这一复合概念统摄二者,旨在整体性地把握 AI 时代知识从创造到传播、从吸收到再生的全过程。进一步看,从口语传播时代到互联网时代,知识

生产与知识流动不断交互作用，并与不同的现实相结合，进而源源不断地为人类文明的进步提供动力。

1. 口语传播时代。在口语传播时代，知识高度依赖人脑的记忆与口耳相传。其内容多为源于狩猎、采集、祭祀等具体生活场景的经验性与地方性知识，常以神话、传说、歌谣等形式承载。部落中的长者、巫师与工匠是主要的知识生产者。由于此阶段的知识流动“局限于身体在场的范围”，^①传播场域极为有限。这使得知识呈现出强烈的“情境化”与“个人化”色彩，导致知识进步异常缓慢，且在口耳相传的链条中极易出现失真、变形与丢失。

2. 文字与手稿时代。文字的发明是人类知识史上的一次重要突破，它使知识得以脱离人脑，实现稳定可靠的外部存储。由此，知识生产开始走向系统化与抽象化，并催生了哲学、法律、天文学、数学等学科雏形。知识生产的主体也转变为祭司、文吏等精英阶层，其载体则是泥板、竹简与羊皮卷。尽管知识的流动范围因商队与信使而得以跨越地理限制，但高昂的成本与复制的困难，导致知识权力被少数主体垄断，形成了集中的知识中心。尽管如此，一个更为深远的历史进程已不可逆转，即在文字的推动下，人类建构知识体系的速度极大地加快了。

3. 印刷时代。印刷术的发明实现了知识的机械化复制，其通过文本标准化、知识保存与传播革新重构社会认知体系。这一阶段知识生产的速度和规模空前提高。报纸、期刊、书籍成为新载体，催生了科学革命和启蒙运动，促进了科学共同体的形成和科学方法的发展。知识的生产者从宗教精英扩展到科学家、学者、作家等世俗知识分子。知识流动也得以大规模、低成本地传播，打破了教会和贵族的知识垄断，促进了民族语言和统一文化的形成，公众阅读能力得到提升，知识的流动从“中心到中心”开始转向“中心到边缘”。

4. 电子传播时代。电报、广播、电视等技术的出现，实现了信息的即时与远程传输，使得知识在社会中以多元载体的形式广泛弥散。在形态上，媒介形态的变化深刻改变了知识分子群体的知识传播方式。^②知识不再局限于文字，而是扩展为包含声音、动态图像在内的多元表达形式，催生了如新闻直播、纪录片、电视讲座等新知识载体。在机制上，知识生产逐渐由专业大众媒体机构主导，流动则呈现出鲜明的瞬时性、同步性与广覆盖性。此阶段的流动模式主要为一对多的广播式传播，大众多数情况下处于被动接收端，主流媒体凭借其渠道垄断与内容掌控，建立起强大的话语体系。

5. 互联网时代。个人电脑和互联网的普及，使得知识实现了数字化、网络化和交互化。知识生产权彻底下放，“全民创作”时代到来。互联网让每个人都成为潜在的知识生产者。知识呈现爆炸式增长，但也带来了信息过载和真假难辨的问题。知识流动模式则变为多对多的网状结构，知识可以在全球范围内实时、互动地流动。搜索引擎和社交算法成为知识分发的主要渠道，但数字技术的大规模运用也让人类遭遇“信息乌托邦”，受困于“信息茧房”。

（二）AI 时代知识生产流动的特殊性

从传统知识生产模式到交叉学科研究的新模式，当代社会知识的生产方式一直在发生变化。^③而随着新一代人工智能介入知识生产，由人类主导的知识生产范式面临再一次革命性变革，知识流动也随之迭代发展（表1）。此时机器已成为能自主或半自主生成、关联与优化知识的“认知伙伴”，知识的生产流动由此呈现出五大特殊性。

第一，生产主体的颠覆，即从“人类独白”到“人机协智”。此前的所有技术，从笔墨到印刷机再到计算机，本质仍是人类心智的被动工具，它们无法理解内容，亦无法创造新知。而今，

① 邱泽奇：《人机互生：范式革命与知识生产重构》，《探索与争鸣》2024年第1期。

② 参见张振宇：《知识分子与电视：角色变迁、功能转型及其反思1988—2011》，武汉：华中师范大学出版社，2015年，第2页。

③ 参见迈克尔·吉本斯等：《知识生产的新模式：当代社会科学与研究的新模式》，陈洪捷、沈文钦等译，北京：北京大学出版社，2011年。

表 1 知识生产流动演进历程

时代	核心技术	知识生产特征	知识流动特征
口语时代	语言	经验化、个人化	范围小、易失真
文字手稿时代	文字	系统化、精英化	成本高、被垄断
印刷时代	印刷术	规模化、标准化	大规模、低成本
广电时代	广播电视	机构化、多媒体化	一对多、瞬时性
互联网时代	互联网	民主化、爆炸式	多对多、交互性
AI 时代	人工智能	自动化、生成式、关联化	个性化、智能化、人机协同

① 参见殷杰：《生成式人工智能的主体性问题》，《中国社会科学》2024 年第 8 期。

AI 已能作为协作者甚至创作者深度介入，其通过更敏锐的感知能力、更强大的推理能力和更适切的行动能力，正展现出一种新型主体性，^①知识生产也随之从纯粹的人类心智活动，变成了人机交互的对话过程。第二，生产模式的革命，即从“推理分析”到“关联涌现”。传统知识生产倚重人类的逻辑推理、实验归纳和理论建构，旨在基于已知原理去探索未知。而 AI 大模型擅长在海量数据中寻找非显而易见的关联和模式，它生产知识的方式是“涌现”——从统计概率中浮现出新的见解、答案甚至创作风格，其内在逻辑有时对人类而言是“黑箱”。第三，载体形态的演变，即从“可读化”到“可计算化”。在过去，知识主要被编码为文字、图表、公式等人类可读的形式，其价值的释放必须经由人类的阅读与理解。而在 AI 时代，知识则被深度嵌入到模型的参数与向量之中，其形态本身即是机器可理解、可操作的。这也体现了知识的运行逻辑，从必须跨越人类的“理解门槛”，转向了遵从机器的“计算逻辑”。第四，流动方式的升级，即从“搜索—分发”到“解读—对话”。过往知识流动的范式是检索和分发，核心是“找到信息”和“推送信息”。而在 AI 时代，用户可以直接与内化了知识的智能体进行多轮对话，通过指令让其执行总结、解释乃至创作等任务。这意味着，知识获取的核心诉求已从高效地“寻找信息”升级为深入地“理解洞察”。第五，循环速度的极致，即从“代际积累”到“实时迭代”。在过去，知识的创造、验证到应用要经历一个漫长的周期，而 AI 时代知识生产流动的闭环被极大地压缩和加速。AI 能够实时吸收新数据，并瞬时将其整合进自身的知识体系，通过交互接口实现全球同步。与此同时，持续的人机互动构成了一个高效的反馈回路，不断驱动系统进行自我优化与升级。

概言之，AI 时代的智能体不再是被动工具，而是成为一个能够与人类共同思考、共同创造的交互生产者和协同流动者。从生产端来看，知识的生产者不再仅仅是人类，AI 成为强大的协作者甚至独立创作者；从流动端来看，知识的流动从“人找信息”演变为“信息找人”，AI 推荐系统极大地提升了知识分发的效率，“人机协智”成为一种全新的知识流动模式。

AI 时代知识生产流动的失序风险

② 尼尔·波斯曼：《技术垄断：文化向技术投降》，何道宽译，北京：中信出版集团，2019 年，第 3 页。

美国学者尼尔·波斯曼指出：“每一种技术都既是包袱又是恩赐，不是非此即彼的结果，而是利弊同在的产物。”^②AI 时代的技术突破为知识生产流动带来前所未有的赋能的同时，也引发了一系列失序风险。具体如下：

其一，真实性风险，表现为知识生态的“污染”与信任崩塌。AI 时代知识生态的污染首先源于 AI 的“幻觉”，即模型会生成看似合理实则虚构的内容，这些内容与错误信息、偏见数据混

杂在一起，构成了大规模、高逼真的“信息迷雾”。^①其中存在着一种“奇幻社会”的危机，AI 幻觉与后真相相互递归，使得虚构内容能够自我强化、循环验证，最终模糊甚至取代客观实在。其次，更为严峻的是“AI 污染”问题，即 AI 生成的虚假内容被大量发布到互联网，并作为训练数据反哺新一代 AI 模型，导致错误如病毒般在代际模型间复制与放大，造成“能指变异”与“所指危机”，^②使语言符号与其指代的真实世界脱钩。这种系统性失真最终将侵蚀社会信任的基石，公众无法区分 AI 生成内容的真伪，容易陷入“什么都不信”的普遍怀疑主义，科学共识、新闻权威乃至历史事实的确定性都将被动摇。

其二，公平性风险，表现为知识获取的新鸿沟与认知不平等。“智能鸿沟”超越了传统“数字鸿沟”对设备与网络接入的关注，演化为在资源、能力与收益上的差异，进而引发社会不平等和不公正。具体而言，首先是资源鸿沟。顶尖 AI 模型的训练与使用成本高昂，并优先服务于富裕国家、大型企业与高收入群体，而发展中国家与弱势群体则被边缘化，知识获取的“马太效应”空前加剧。其次是能力鸿沟，即使能够接触 AI，个体在算法素养、提示词工程与批判性使用能力上的差异，决定了其能从 AI 中获取知识价值的效率与深度，导致认知结果的不平等从“知识沟”深化为“实践沟”。^③最终，这演变为权力与收益的鸿沟，掌握 AI 资源的阶层能更高效地生产知识、创造财富，从而巩固其优势地位；而缺乏资源的群体则在认知和经济上双重落伍。此外，政务、教育、医疗等关键领域的 AI 应用本应普惠大众，但若使用不当，反而会制度化地再生产并固化一系列新的社会不平等，引发“模型输出对弱势群体隐性排斥加剧的结果鸿沟”。^④

其三，主体性风险，表现为人类认知外包与批判性思维退化。AI 时代知识流动的便捷性，潜藏着侵蚀人类认知主体性的深刻风险。换言之，技术的自决性以及自主知识的“编造”过程正在逐渐取代人的主体性思维活动。当生成式 AI 能够即时提供答案、撰写文案，甚至进行逻辑推理时，人类倾向于将越来越多的认知任务“外包”给机器。这种便利性无形中导致“认知惰性”的滋生，使个体主动思考、深度记忆和独立判断的意愿与能力不断退化。批判性思维这一人类理性核心的锤炼，依赖对信息的主动甄别、逻辑链条的细致梳理以及不同观点的反复衡量。然而，在 AI 提供“一站式”答案的模式下，这些思维过程被大幅压缩甚至绕过。长此以往，人类的认知能力可能发生结构性改变，正如工具塑造了我们的手掌，过度依赖 AI 也可能重塑我们的大脑，使其在与机器的“共生”中逐渐丧失自主探索与质疑的活力。更深远风险在于，我们可能将 AI 基于数据关联得出的概率性结果误当作绝对真理，放弃了对知识本质的追问和对世界本源的批判性洞察，最终从知识创造与批判的主体，异化为 AI 知识输出的被动接受者。

其四，权力性风险，表现为知识权力的垄断与新型数字霸权。AI 时代的知识生产流动，正催生一种由技术和资本共同主导的新型权力结构，其中潜藏着知识垄断与数字霸权的风险。这种权力风险首先体现在技术层面，全球顶尖的生成式 AI 模型几乎由少数几家西方科技巨头开发和掌控，它们通过垄断庞大的高质量数据集、顶尖的算力基础设施以及核心的算法模型，实质上把持了全球知识生产与分发的源头。其次是文化与语言层面的风险。当前大模型存在中心语言与边缘语言的不平等结构，以英语为代表的中心语言所占语料呈现结构性支配地位，其内在的价值观、思维模式和文化偏好具有强烈的西方中心主义色彩。当这些模型成为全球知识流动的主要中介时，非西方世界的语言、本土知识与文化观念面临被边缘化、扭曲甚至消解的危险，造成全球文化多样性的“静默危机”。最终，这种技术霸权与文化霸权将转化为规则制定权，巨头们通过其平台规则和算法黑箱，无形中决定了哪些知识被看见、被传播以及如何被诠释，从而塑造公共议程与社会共识。

① 张梦晗、沈文乾：《错位的升级与结构性失序：AI 幻觉主导的信息迷雾风险与分类治理可能》，《传媒观察》2025 年第 10 期。

② 赵靓：《AI“污染”：数智时代的能指变异与所指危机》，《民族学刊》2025 年第 8 期。

③ 虞鑫、王德龙：《从知识沟到实践沟：算法素养与短视频内容获取的不平等研究》，《新闻与写作》2025 年第 9 期。

④ 张会平、马明振、钟毅：《政务大模型应用的智能鸿沟与风险应对》，《重庆社会科学》2025 年第 6 期。



其五，文明性风险，表现为知识传承的“断裂”与历史虚无主义。AI的介入可能对人类社会长期形成的知识传承机制构成根本性冲击，引发文明延续的“断裂”风险。一方面，AI的知识生成严重依赖训练数据，这导致其不可避免地带有“数字偏见”，即过度关注数字时代活跃的、主流的观点，而历史上大量未被数字化、非主流的、口传的或物质性的知识则面临被系统性遗忘的危机。这种数据驱动的“选择性记忆”正在塑造一个扭曲的、扁平化的历史镜像。另一方面，当AI能够轻易地生成、混仿甚至“修正”历史叙事与文化遗产时，历史的客观性与严肃性会受到挑战。它可能基于特定模式“缝合”出看似合理实则虚构的历史记载，或为历史虚无主义提供强大的技术工具，使篡改历史、解构经典变得前所未有的容易和逼真。长此以往，承载民族认同与文化基因的集体记忆可能变得支离破碎，文明赖以延续的连贯性、独特性和深层价值在代际传递中出现断层，最终动摇我们理解自身由来与未来走向的文化根基。

总体而言，人工智能时代的知识失序风险，不再是简单的“信息过载”或“假新闻”问题，而是一场系统性、结构性的认知危机。它挑战了真实性、公平性、人类主体性、权力平衡乃至文明延续的根基。这正是为什么笔者提出“负责任”的知识生产流动这一治理目标，并寻求“敏捷—包容”转向的原因所在。

知识生产流动的“敏捷—包容”混合治理

从某种程度上说，知识失序风险本质上是工业时代形成的线性、静态、封闭的流动生产逻辑，与AI时代非线性、动态、开放的知识生态之间结构性错位的集中体现。要破解这一困局，需以“负责任”为价值基点，实现敏捷机制对包容性的深度嵌入和包容机制为敏捷性的有效赋能，从而构建一个既能适应技术迭代速度，又能承载社会公共价值的“敏捷—包容”混合治理体系。

（一）“负责任”的知识生产流动：“敏捷—包容”转向

面对AI驱动的知识生态复杂性，既有的几种治理模式均显露出其内在的、难以自洽的局限性，陷入系统性失灵的困境。具体如下：

其一，传统刚性监管之困。这种以政府为中心、依赖行政命令与立法手段的“命令—控制”型模式，其核心缺陷在于治理速度与技术迭代速度之间的“速度鸿沟”。法规的制定、修订与执行周期漫长，难以匹配AI技术的指数级演进。其后果是双重的：一方面，滞后的监管无法有效预见和防范如“AI污染”、算法偏见等新型系统性风险，导致治理失效；另一方面，过于僵化与宽泛的规则可能扼杀技术创新的多样性与不确定性，阻碍知识生产潜能的释放。

其二，纯粹技术自治之谬。与刚性监管相对的另一端，是主张完全由市场与科技企业通过代码、算法与平台规则进行自我规制的技术自由主义路径。此路径谬误在于，其天真地假设技术可独立于社会价值而中立演进。然而，算法本身即内嵌着设计者的价值判断与文化偏见，缺乏外部制衡与价值引导，最终容易导向由技术精英与资本力量主导的“算法霸权”，加剧知识权力的集中与垄断，侵蚀社会公平与多元性。

其三，片面追求敏捷之弊。若在治理中仅强调响应速度与决策效率，而忽视对包容性价值的审慎考量，那么“敏捷”极易蜕变为技术精英“闭门造车”的合理化工具。这种缺乏广泛社会共识与多元利益协商的快速决策，虽然可能在短期内解决技术性问题，但从长远看，会因忽视公平、正义等社会核心关切而丧失合法性基础，并可能固化甚至加剧现有的“智能鸿沟”，引发更深层

的社会信任危机与撕裂。

其四，片面追求包容之限。反之，若在设计、治理规则时，片面追求所有利益相关者的全面参与和共识的绝对达成，则可能导致协商流程冗长，决策成本高昂、议而不决等问题。在技术与社会情境快速演进的背景下，这种对“完美包容”的过度追求，将使治理行动严重滞后，错失干预与引导的关键时间窗口，最终导致包容本身因无法产生实效而失去意义。

上述四种单一治理模式的失灵揭示了一个核心命题，即应对 AI 时代知识治理的复杂性，需要摒弃非此即彼的二元对立思维，转而寻求一种能够有机融合效率与公平、创新与规范、灵活与稳定的更高阶治理范式。鉴于此，本文提出“敏捷—包容”的混合协同治理范式，其内在逻辑体现为一种相互依存、相互塑造的协同关系：二者是应对技术复杂性与价值复杂性的“一体两面”，共同构成“负责任”的知识生产流动的治理双轨。一方面，敏捷为包容赋能。敏捷治理为包容性参与构建了高效的实施载体，促进知识生产与流动的效率。它使政府、产业、学界等多元主体的知识、意见与诉求，能够以更低成本、更高频、更灵活的方式持续纳入治理流程，激发创新网络主体之间的显性和隐性技术、信息等知识的扩散和溢出。^①这实质上解决了传统包容性治理效率低下的痛点，让广泛深度的社会参与变得可持续、可操作，从而实现了包容性的过程化与常态化。另一方面，包容为敏捷导航。包容性治理为快速迭代的敏捷进程设定了不可或缺的价值基石与规范框架。它通过建立多元主体的实质性参与机制，确保在决策的全生命周期中，能够持续注入公共利益、社会伦理、文化多样性、个体尊严等规范性考量。这确保了治理的敏捷响应不会仅在技术效率的单一轨道上滑行，从而防止其偏离公共福祉的轨道，增强整个治理体系应对外部冲击的韧性与可持续性。

① 参见郑展：《知识流动与区域创新网络》，北京：中国经济出版社，2010 年，第 9 页。

（二）“敏捷—包容”治理视角下知识生产流动的“四用”及其要素

前述“敏捷—包容”的治理范式为解决 AI 时代知识失序风险提供了方法论基础，而要将这一理念转化为具体实践，则需进一步析出可操作的治理要素。基于此，本文提出以“知识可用—知识有用—知识易用—知识善用”为核心的多维治理体系及其要素。该体系将治理原则具象化为四个递进层次：从确保知识“有无”的生态基础，到衡量知识“好坏”的质量内核，再到评估知识流动“难易”的交互体验，最终达至知识善用“对错”的伦理价值。

1. 知识可用：生态基础层的“有无”问题

知识可用是指知识在物理、数字和法律层面具备可获取性与可及性，是整个知识生态良性循环的基石。在 AI 时代，这不仅意味着知识资源的物理化存储，更需要构建稳定、可靠的数智化知识基础设施。知识可用的治理目标旨在建立一个可及、可得、可获取的知识供给体系，确保社会拥有源源不断的“知识泉眼”。其涉及如下要素：

（1）知识的可及性。知识应是可达、可得、可获取的，包括物理和数字层面的访问便利。在 AI 时代，这意味着需要保障基础网络覆盖、算力资源公平分配以及知识平台的稳定运行。治理应重点关注数字基础设施建设，消除地域和经济的接入壁垒，确保任何个体都能获得所需的知识资源，为知识社会的构建奠定坚实基础。

（2）知识的平等性。它要求治理措施必须致力于消弭“智能鸿沟”，关注资源、能力和收益的公平分配，确保知识红利不被特定阶层垄断。这一原则要求通过政策干预和技术创新，为弱势群体提供补偿性的知识接入支持，防止知识权力过度集中，维护知识生态的多元与平衡。

（3）知识的权属性。知识是有所有权的，涉及如版权、专利等。在 AI 生成内容激增的背景



下,权属性要求平衡知识产权保护与知识共享之间的关系。通过建立适应 AI 时代的知识产权制度,明确人类创作者与 AI 系统的权利归属,既保障创新激励,又促进知识的合法流通与再利用,为知识生产提供可持续的制度保障。

(4) 知识的开放性。知识应以开放格式提供,便于传播和再利用。开放知识不仅意味着技术格式的标准化与互操作性,更要求建立合理的开放许可机制。治理应推动公共领域知识的开放获取,鼓励科研数据和政府信息的开放共享,同时建立安全可控的开放边界,防止重要知识资源的不当扩散,实现开放与安全的动态平衡。

简言之,知识可用是 AI 时代知识生产流动的生态基石,其中可及性、平等性、权属性与开放性的协同演进,共同构筑了知识生态的底层架构。这不仅要求我们突破传统物理存储的局限,更需要构建兼具稳定性与适应性的数智基础设施。知识可用性的治理本质上是开放与保护、普及与专属、共享与确权之间持续调适的动态过程。其深层意义在于,通过制度创新与技术革新的双轮驱动,重塑知识的供给模式,使知识资源从稀缺走向丰裕、从垄断走向共享,最终为构建“敏捷—包容”的知识社会生态提供支撑。

2. 知识有用:质量内核层的“好坏”问题

知识有用强调知识内容本身的价值,要求知识不仅真实可信,而且与用户需求高度契合。其强调知识对于使用者具有实用价值和情境相关性,超越了简单的“存在”,转向知识是否能有效满足个体或社会的特定需求。在 AI 时代,知识有用性面临新的挑战:海量信息中的价值筛选、个性化需求的精准匹配、生成内容的可靠性验证。因此,知识有用的治理目标在于确保流动的知识是高质量、高相关、可理解的,能够有效解决实际问题,赋能决策与创新。其涉及如下要素:

(1) 真实性。这是“知识有用”的前提,无法确认真伪的知识不具备“有用”价值。真实性原则致力于保障知识来源的可溯性与内容的可信度,其核心在于对抗“AI 幻觉”与系统性信息污染,以维护社会集体认知所依赖的客观基础。治理需要通过多层次验证机制,包括数据溯源技术、多方事实核查以及可信度评级系统,构建从生产到流动的全链条真实性保障体系,维护社会认知的客观基础。

(2) 可信性。可信性是真实性在认知层面的延伸,它不仅要确保知识本身的真实,更要建立用户对知识系统的持久信任。可信性治理聚焦构建知识来源的可靠性证明机制,包括开发透明的溯源技术、建立内容可信度评估体系,以及推动算法决策的可解释性。通过数字水印、机构认证和可信传播路径追踪等技术手段,使知识生产流动过程可审计、可验证,从而在快速迭代的信息环境中维持稳定的认知信任基础,这是知识从短期真实走向长期可信的关键保障。

(3) 准确性。知识内容应精确无误,正确反映客观事实。在 AI 时代,准确性治理需要应对大规模生成内容的质量控制挑战。这要求建立精确的知识验证体系,包括多重交叉检验、人机交互检验等。同时,通过持续的模型优化和反馈循环,提升 AI 系统的事实一致性,确保知识输出不仅语义通顺,更与客观现实高度吻合。

(4) 时效性。知识应与其所处的时间背景高度相关,及时更新,反映最新进展。在快速变化的 AI 时代,时效性治理需要建立动态的知识更新机制,包括实时数据流处理、版本控制系统和过时内容标识。治理应推动知识库的持续维护,发展能够自动追踪知识演变的智能系统,确保用户获取的知识始终具有当下相关性。

(5) 相关性。知识应与用户的具体场景、任务和认知水平高度适配。相关性要求知识服务系

统能够深度理解用户意图,实现精准的知识匹配。治理应鼓励开发情境感知的 AI 系统,推动个性化推荐技术的发展,同时防止过度个性化导致的“信息茧房”,保持知识生态的多样性与开放性,实现从“泛在知识”到“精准价值”的转变。

(6) 完整性。知识应提供全面信息,避免碎片化。完整性治理要求知识呈现保持必要的上下文和逻辑关联,防止片面解读。在 AI 时代,这需要开发能够自动识别知识盲区、补充背景信息的智能系统,同时建立知识结构化标准,促进系统化知识的构建与传播,确保用户获得立体、多维的认知图景。

由此可见,知识有用是 AI 时代知识生产流动的质量基石。它超越了信息的表层存在,直指知识赋能的核心——通过构建真实性、可信性、准确性、时效性、相关性与完整性的多维质量体系,为快速迭代的知识生态确立稳固的质量内核。这不仅关乎技术层面的精准匹配,更在深层次上塑造了人机协同的信任基础与可信 AI 生态体系,确保了知识从“数据洪流”向“决策智慧”的有效转化。

3. 知识易用:流动交互层的“难易”问题

“易用”原则致力于降低知识生产和流动过程中的门槛与成本,确保不同背景、不同能力的个体和群体都能平等、便捷地从中受益。在 AI 时代,易用性不仅关乎界面友好,更涉及认知负荷的降低和交互效率的提升。其治理目标在于,实现知识生产与流动的低门槛、低成本、无歧视,促进知识社会的包容性发展。知识易用涉及如下要素:

(1) 可检索性。知识应能够被快速、方便地查找,这强调搜索和匹配过程的效率与效果。在 AI 时代,可检索性治理需要优化搜索引擎算法,发展语义理解和多模态检索技术,使知识查找从关键词匹配升级为意图理解。同时,建立统一的知识标识和元数据标准,促进跨平台的知识发现,确保用户在信息海洋中能精准定位所需内容。

(2) 友好性。知识呈现方式应用户友好,如直观的界面或语言。友好性治理要求知识交互系统符合人类认知习惯,提供清晰的信息架构和自然的交互流程。这包括开发适应不同用户群体的界面设计规范,推动通俗化表达和可视化呈现,降低技术使用门槛,使知识获取过程顺畅无碍,提升整体用户体验。

(3) 可操作性。知识应易于转化为实际行动或决策指南。可操作性治理强调知识的实用转化,要求知识内容具备明确的行动指引和价值落地路径。通过开发知识到行动的转化工具,建立最佳实践库和案例指导系统,帮助用户将抽象知识转化为具体解决方案,真正发挥知识的实践价值,弥合知与行之间的鸿沟。

(4) 可访问性。知识应适应不同用户群体能力,符合无障碍设计。可访问性治理要求知识服务体系充分考虑各类用户需求,包括残障人士、数字技能薄弱群体等。通过推行普适设计原则,开发多模态交互方式,提供个性化适配功能,确保知识生态的包容性,让每个人都能平等地参与知识社会建设。

因此,知识易用是 AI 时代知识生产流动的效能枢纽,旨在减少障碍,保持知识流动的畅通无阻,^①进而提升知识从潜在价值到实际效能的转化效率。在 AI 重塑知识交互模式的当下,易用性通过构建人机协同的无缝接口,大幅降低了知识获取的认知负荷与技术门槛,也通过优化交互体验使知识资源在不同群体间高效流动与深度转化。其深层意义在于,通过提升知识分发的精准度与普惠性,知识易用性有效激活了知识生态的内在活力,弥合“智能鸿沟”,为构建真正包容

① 参见吴建中:《知识是流动的》,上海:上海远东出版社,2015年,第45页。



① 郑永年：《AI时代的中国知识重建》，香港：三联书店（香港）有限公司，2025年，第65页。

高效的知识社会提供赋能支点。

4. 知识善用：伦理价值层的“对错”问题

“知识的生产有其特殊性，需要一个特殊的思考环境”，但如今许多制度“反而成了知识生产的严重阻力”。^①因此，“知识善用”作为“四用”框架中的最高层次，旨在通过制度规制与环境建设规定知识应用的伦理方向和长远责任，强调知识流动的最终目的是增进人类福祉与促进文明可持续发展。在AI时代，善用治理需要应对技术赋能带来的伦理挑战，确保知识创新服务于人类共同利益。知识善用的治理目标是确保知识生产流动符合人类共同伦理，服务于公共利益和文明的长远延续，实现科技向善。其涉及如下要素：

（1）人本性。坚守人的主体性地位，确保技术服务于人的认知发展与整体福祉。这是“善用”的终极价值归宿。一切知识应用都必须以促进人的全面发展、维护人的尊严和主体性为最高准则，防止技术异化。治理应确保AI系统始终处于人类控制之下，知识创新以增强人类能力为目标，防止人在技术应用中沦为客体或被边缘化。

（2）韧性。培育知识生态系统自身抵御污染、对抗攻击，并从干扰中恢复的能力，确保其长期健康与稳定。一个脆弱的知识生态系统无法持续提供可靠的知识。治理需构建能够抵御攻击、能从污染中快速恢复的系统，通过多元化知识来源、冗余设计和自适应机制，增强知识生态的抗冲击能力，保障知识服务的连续性与稳定性。

（3）伦理性。知识创造与使用应符合道德规范，避免滥用或伤害。伦理性治理要求建立全面的AI伦理框架，涵盖隐私保护、公平正义、知情同意等核心价值。通过伦理审查等机制的贯彻和普及，确保知识活动始终在道德边界内进行。

（4）可持续性。知识本身及其应用应能支持并促进社会、环境与经济的长期繁荣与良性发展。知识的生产与应用应有利于人与自然的长期和谐共生。治理需引导AI知识流向气候变化、生物多样性保护、绿色能源等可持续发展目标，建立知识创新的环境影响评估机制，确保技术进步与生态保护协同共进。

（5）向善性。知识旨在促进个人福祉与社会进步，引导产生积极影响。向善性治理强调知识应用的价值导向，要求技术发展服务于社会公益和人类福祉。通过建立向善创新激励机制、负责任的文化和社会价值评估体系，引导知识资源投向教育、医疗、减贫等关键领域，最大化知识的正向社会效益。

简言之，人本性、韧性、伦理性、可持续性与向善性构成的“知识善用”治理维度，回应了AI时代知识生产流动的特殊性与制度阻力的现实困境，其以“人”为核心锚点，规范知识的生产逻辑、流动路径与应用边界，让技术创新始终扎根于人类共同利益的土壤，既破解知识生产流动的失序风险，又激活知识服务文明进步的深层潜能，推动人类在智能时代实现更高质量的知识创造与文明演进。

综上，通过系统拆解“敏捷—包容”治理视角下知识生产流动“四用”体系的层级逻辑、治理目标及核心要素，本文将“四用”及其核心要素梳理汇总如表2所示，这为后续构建AI时代“负责任”的知识生产流动“四用”框架奠定了基础。

（三）AI时代“负责任”的知识生产流动“四用”框架构建

根据前文“敏捷—包容”治理视角下知识生产流动的“四用”及其要素剖析，本文构建起一个以“负责任”为内核、“敏捷—包容”为治理底色的“知识四用”框架。该框架并非孤立要素

表 2 “敏捷—包容”治理视角下知识生产流动的“四用”及其要素

知识四用	知识可用	可及性：知识是可达、可得、可获取的，包括物理和数字层面的访问便利
		平等性：知识面前人人平等，不受地域、经济或社会背景的限制
		权属性：知识是有所有权的，涉及法律保护，如版权、专利等
		开放性：知识应以开放格式提供，便于传播和再利用
	知识有用	真实性：知识基于事实和证据，避免虚假或误导
		可信性：知识来源的可溯性与内容的可信度，对抗“AI 幻觉”与信息污染，维护社会认知的客观基础
		准确性：知识内容精确无误，正确反映客观事实
		时效性：知识与其所处的时间背景高度相关，及时更新，反映最新进展
		相关性：知识与用户需求高度匹配，具有实用价值
		完整性：知识提供全面信息，避免碎片化
	知识易用	可检索性：知识能够被快速、方便地查找到，强调搜索和匹配过程的效率与效果
		友好性：知识呈现方式用户友好，如直观的界面或语言
		可操作性：知识易于转化为实际行动或决策指南
		可访问性：知识适应不同用户群体能力，符合无障碍设计
	知识善用	人本性：坚守人的主体性地位，确保技术服务于人的认知发展与整体福祉
		韧性：培育知识生态系统自身抵御污染、对抗攻击、并从干扰中恢复的能力，确保其长期健康与稳定
		伦理性：知识创造与使用符合道德规范，避免滥用或伤害
		可持续性：知识本身及其应用应能支持并促进社会、环境与经济的长期繁荣与良性发展
		向善性：知识旨在促进个人福祉与社会进步

的简单叠加，而是一套层层递进、动态耦合的治理体系，力图实现知识生态的效率与伦理共生。

首先，“知识可用”作为框架的生态基础层，为整个知识生产流动体系筑牢根基。这一层面以消除知识获取壁垒为核心目标，通过可及性、平等性、权属性与开放性的协同推进，应对公平性风险中的资源鸿沟与权力垄断隐患。在敏捷治理维度，需展开动态资源调配，针对 AI 技术迭代带来的算力需求变化，实时优化数字基础设施布局，确保知识供给与技术演进同频。在包容治理维度，需将平等性嵌入制度设计，一方面通过公共政策倾斜，为发展中国家、低收入群体提供普惠性算力与知识平台服务；另一方面探索构建适应 AI 时代的知识产权“前沿”制度，在保护创作者权益的同时，明确 AI 生成内容的权属边界，并保障公共领域知识与数据的合法共享，防止知识资源被少数主体垄断。这一层面的核心价值在于让知识实现“全民可及”，确保知识生态的基础普惠性与制度适应性，为后续知识的质量提升、效能转化与伦理实践奠定基础。

其次，“知识有用”作为框架的质量内核层，为整个知识生产流动体系确立质量基准，回应知识的真实性风险与认知偏差问题。其以精准赋能为目标，通过真实性、可信性、准确性等要素的系统构建，确保知识流动能够真正服务于个体决策与社会发展。在敏捷治理层面，需实现实时动态的知识验证与更新，利用数智化技术实现知识溯源，通过人机协同核查系统快速识别并标注 AI 幻觉、虚假信息，同时开展知识版本迭代管理，确保时效性与完整性。在包容治理层面，强调知识价值的多元适配，既通过个性化推荐技术满足不同用户的场景化需求，又警惕“信息茧房”的形成，通过算法干预保障知识供给的多样性与全面性，让不同认知水平、不同需求场景的用户都能获取适配的高质量知识。这一层面的关键意义在于，为整个知识生态树立可信赖的质量规范，将知识自身提纯为一种“可靠资产”，而非“其他类型资产的附属物”，^①进而重建知识生态的信任基础。

再次，“知识易用”作为框架的流动交互层，是连接知识供给与需求的效能枢纽，着力破解

① 马克斯·博伊索特：《知识资产——在信息经济中赢得竞争优势》，张群群、陈北译，上海：上海人民出版社，2021 年，第 2 页。

① 吴建中：《知识是流动的》，上海：上海远东出版社，2015年，第3页。

② 参见安小米、高俊杰：《数字政府建设中数据开放的标准化协同：要素识别与框架构建》，《中国图书馆学报》，2025年第6期。

数字鸿沟向智能鸿沟演化等风险。其以低门槛、高效率为导向，通过可检索性、友好性、可操作性、可访问性等要素的优化组合，降低知识生产与使用的技术门槛和认知成本。在敏捷治理维度，需推动交互界面与知识获取流程的持续优化与快速迭代，使知识从各自分散独立的状态下释放出来并处于可检索、可获取、可利用的有序状态，为后续高效地开放和共享知识创造条件。^①同时，建立知识交互的快速反馈渠道，根据用户行为数据实时优化界面设计与操作流程。在包容治理维度，坚持普适设计原则，一方面针对老年人、残障人士等群体开发无障碍交互功能，如语音导航、大字模式、手语翻译等；另一方面通过知识通俗化转化、可视化呈现等方式，降低数字技能薄弱群体的认知负荷，让知识流动突破能力壁垒，实现知识面前“人人皆可参与、人人皆能受益”。

最后，“知识善用”作为框架的伦理价值层，是“负责任”目标的最终归宿，直面知识的主体性、权力性与文明性风险。其基于以人为本、文明延续的核心导向，通过人本性、韧性、伦理性、可持续性、向善性等五要素的价值锚定，确保知识创新始终服务于人类共同福祉。在敏捷治理层面，需将上述伦理价值前瞻性地融入技术标准与政策制定过程，系统性地前置风险研判与规避机制；同时面对知识权力集中、历史叙事扭曲等问题，应确保在算法出现偏见或扭曲时，能够快速启动人类主体的干预措施，维护人类在知识生产与流动中的主体地位。在包容治理层面，构建多元利益相关方的标准化协同体系，通过吸纳政府、产业界、学术界、公民等各方力量参与知识伦理规则制定，实现从无序到有序、从混乱到清晰、从复杂到简单、从散乱到统一的标准化协同效益。^②既尊重不同文化背景下的价值差异，又坚守人类共同伦理底线，确保知识应用在不偏离公共利益轨道的同时，维护文化多样性与文明连续性。

简言之，“知识四用”框架构成了一个“基础—核心—保障—归宿”的梯次递进、相互支撑的体系结构。“可用”解决知识“能否获取”的前提问题，“有用”解决知识“是否可靠”的核心问题，“易用”解决知识“能否用好”的效能问题，“善用”解决知识“应该如何用”的伦理问题。而“敏捷—包容”的治理理念贯穿始终，使框架既具备应对技术迭代的动态适应性，又拥有承载多元价值的社会包容性。该框架通过各层面要素的协同联动，能够系统性回应 AI 时代知识生产流动的真实性、公平性、主体性、权力性、文明性等多重风险，将“负责任”从抽象理念转化为可操作、可落地的治理实践，为构建人机共生背景下的健康知识生态提供清晰的行动蓝图（图 1）。

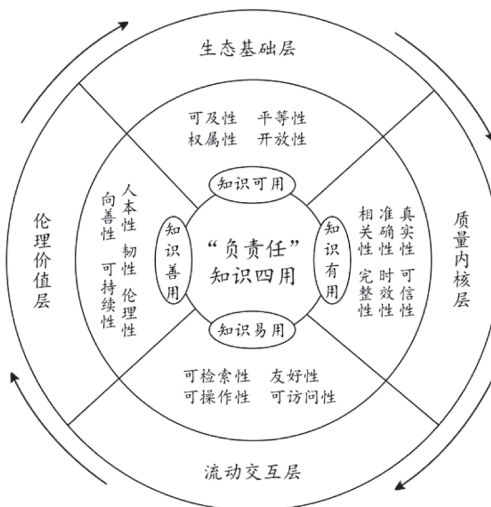


图 1 “负责任”的知识生产流动“四用”框架

结语：在流动的世界中寻求责任

罗素曾指出：“人类的全部知识都是不确定的、不准确的和片面性的。”^①在 AI 驱动下，知识以前所未有的速度、规模和复杂性流动的方式呈现，不确定性被急剧放大，甚至构成严峻的认知挑战。然而，我们不能因噎废食，更不能退回封闭与僵化的治理老路，而是需要在这充满不确定性的流动世界中，为知识寻求一份“负责任”的确定性。本文正是对这一时代命题的系统性回应尝试，主要研究贡献如下：

第一，文章系统揭示了 AI 时代知识生产流动的特殊性及其引致的多维风险。文章通过梳理历史脉络，指出 AI 时代的知识生产流动在主体、模式、形态、方式与速度上发生了根本性变革，机器从被动工具演变为“共生伙伴”。这一变革在带来效率革命的同时，也引发了真实性、公平性、主体性、权力性及文明性五个维度的失序风险。这些风险交织缠绕、相互影响，构成了一个超越传统信息过载结构的全新认知环境危机，凸显了进行有效治理的紧迫性与复杂性。

第二，文章明确提出了“负责任的知识生产流动”这一核心治理目标，并确立了“敏捷—包容”协同的混合治理范式。本文批判性地审视了传统刚性监管、纯粹技术自治以及片面追求敏捷或包容等单一治理模式的局限性，论证了在 AI 时代的超级复杂性面前，任何非此即彼的路径均已失灵。进而，本文主张必须转向一种“敏捷—包容”协同的治理范式，其以“负责任”为价值基点，实现敏捷性与包容性的辩证统一与相互赋能。敏捷治理为包容性参与提供高效载体，使其可持续、可操作；包容治理则为敏捷响应设定价值航向，确保其不偏离公共福祉的轨道。二者共同构成了驾驭知识洪流、应对复杂风险的治理双轨。

第三，文章将抽象的治理理念转化为可操作、可落地的实践框架，构建了“知识可用—知识有用—知识易用—知识善用”的“四用”治理框架。该框架有机融入了“敏捷”与“包容”的治理底色，同时将“负责任”的知识生产流动目标解构为四个梯次递进、相互支撑的治理层次：“可用”是生态基础，解决知识“能否获取”的前提问题，确保知识的可及性、平等性、权属性与开放性；“有用”是质量内核，解决知识“是否可靠”的核心问题，确立真实性、可信性、准确性、时效性、相关性与完整性的质量规范；“易用”是流动交互保障，解决知识“能否用好”的效能问题，通过可检索性、友好性、可操作性、可访问性降低使用门槛；“善用”是伦理价值归宿，解决知识“应该如何用”的根本问题，以人本性、韧性、伦理性、可持续性、向善性锚定知识发展的最终方向。

当然，迈向“负责任”的知识生产流动是一个不断演进的动态迭代和持续发展进程。本文提出的“敏捷—包容”治理范式与“知识四用”框架亦并非一成不变，而是一个开放、动态且有待在实践中持续检验、丰富与迭代的治理基模。在技术与社会相互塑造的进程中，如何进一步细化各治理要素的评估指标、如何在全球范围内构建有效的治理协调机制、如何在保障安全的前提下充分释放知识的创新潜能，均是后续值得深入探讨的方向。

编辑 孙冠豪

① 罗素：《人类的知识：其范围与限度》，张金言译，北京：商务印书馆，1983年，第617页。