

引用格式: 隆云滔, 刘海波, 许哲平, 等. 关于构建我国人工智能开源创新生态体系的建议. 中国科学院院刊, 2025, 40(3): 453-458, doi: 10.16418/j.issn.1000-3045.20250228001.

Long Y T, Liu H B, Xu Z P, et al. Suggestions on building China's artificial intelligence open source innovation ecosystem. Bulletin of Chinese Academy of Sciences, 2025, 40(3): 453-458, doi: 10.16418/j.issn.1000-3045.20250228001. (in Chinese)

关于构建我国人工智能开源创新生态体系的建议

隆云滔^{1,2} 刘海波^{1,2} 许哲平^{3,4,5} 包云岗^{6*} 武延军⁷

1 中国科学院科技战略咨询研究院 北京 100190

2 中国科学院大学 公共政策与管理学院 北京 100190

3 中国科学院文献情报中心 北京 100190

4 学术期刊新型出版与知识服务重点实验室 北京 100190

5 中国科学院大学 经济与管理学院 信息资源管理系 北京 100190

6 中国科学院计算技术研究所 北京 100190

7 中国科学院软件研究所 北京 100190

摘要 随着以DeepSeek为代表的开源大语言模型发布和深入应用, 开源创新在全球人工智能(AI)技术的快速发展中发挥了越来越重要的作用。通过技术创新降低成本、提升性能和应用示范, DeepSeek开辟了一条颠覆性的创新路径, 提高了我国AI的研究应用水平和国际影响力。为了进一步推动我国AI开源创新生态体系的建设, 促进AI技术的持续创新和迭代, 文章分析当前我国AI开源创新生态中存在的突出问题及相关政策的不足, 并从顶层设计、基础设施建设、多元参与主体培育、开源创新运行机制完善等4个维度提出加快构建我国AI开源创新生态体系的建议。

关键词 人工智能, 开源创新, 基础设施, 生态体系建设

DOI 10.16418/j.issn.1000-3045.20250228001

CSTR 32128.14.CASbulletin.20250228001

2025年1月下旬, 杭州深度求索人工智能基础技术研究有限公司成功发布了其自主研发的开源大模型

DeepSeek-R1, 这一突破性成果不仅为人工智能(AI)领域提供了降低成本和提升性能的创新路径, 也成为

*通信作者

资助项目: 国家社会科学青年基金(24CJY018)

修改稿收到日期: 2025年3月10日

我国突破国外技术遏制、提升前沿领域核心竞争力的重要标志，推动我国AI研究水平和应用能力迈上新台阶。尽管DeepSeek引发了全球性的关注，但在AI领域我国的整体实力与美国相比仍存在显著差距。例如，在美国斯坦福大学2024年11月发布的“全球人工智能活力排行榜”（Global AI Vitality Ranking）中，中国虽以40.17分位居第2，但远低于美国的70.06分，尤其在研发投入、人才教育、基础设施等方面与美国的差距明显。

开源创新是AI领域取得当前成就的关键因素之一，美国Meta公司的LlaMA和中国DeepSeek等开源项目的成功再次验证了这一点。因此，加快构建我国AI开源创新生态体系，对于我国抢占AI创新制高点意义重大，未来需要进一步加大对开源创新的支持力度，完善相关政策与基础设施，以推动我国AI创新持续深入发展。

1 我国AI开源创新生态存在的突出问题

1.1 相关政策保障不足

（1）主体政策缺乏“系统整合性”。虽然从国家到地方已经通过顶层设计、专项政策等方式明确了AI产业发展的战略地位，但缺乏AI与开源建设相结合的具体规划，尚未形成“顶层设计—专项政策—具体措施”系统性政策体系。美国2023年发布的《国家人工智能研发战略规划》（*National Artificial Intelligence Research and Development Strategic Plan*）中，明确提出要“开发开源软件库和工具包”^①。英国在2025年1月发布的《人工智能机遇行动计划》（*AI Opportunity Action Plan*），也明确要求“基础设施具有互操作性、代码可重用性和开源性”^②。

（2）关联政策缺乏“积极响应性”。一些政策围绕开源社区、治理规则和标准、人才培养、国内外合作等给出了原则性指导，但缺乏具体规范和细则，产业链和技术链的各相关方没能有效参与，难以为开源创新生态体系构建提供必要支撑。

（3）落地举措缺乏“互动协同性”。例如，现有评价机制更多关注技术性贡献，对过程性等非技术贡献重视不足；激励方式相对单一，企业、科研机构和个人等主体通过开源生态所能获得的资源反哺和产业转化能力相对有限，难以形成有效激励。

1.2 生态稳定性不足

（1）开源生态共生关系先天脆弱。开源天然的“公共属性”与企业固有的“盈利追求”，决定了AI开源创新生态体系建设必然面临利益之争与角色冲突——生态内外部需求的矛盾、多元参与主体的竞合及绩效目标的差异，使得开源创新生态共生关系极易受到改变乃至破坏。AI技术快速演化下的技术及产业需求变化，也会传导并影响生态共生关系，进一步增加不稳定性。

（2）开源要素对外依赖度过高。国内AI开源框架多建立在国外原生框架（如PyTorch、MLIR等）之上，一些关键核心技术仍依赖国外主导的开源项目（如Ollama、Numpy等），常用开源许可证大多来源于美国机构（如Linux基金会、Apache基金会等），国内机构和开发者严重依赖国外代码托管平台及社区（如GitHub、Hugging Face等）。但是，目前Hugging Face在国内已经无法直接访问。而GitHub在国内的访问经常不太稳定，此前还曾对伊朗和叙利亚等国家的开发者进行过限制。多因叠加，致使我国开源生态稳定运行面临较大风险。从技术上看，AI的技术堆栈从大模

① 美国发布《国家人工智能研发战略规划》2023 更新版. (2023-06-12). https://dsj.hainan.gov.cn/zcfg/gwfg/202306/t20230612_3433750.html.

② AI Opportunities Action Plan. (2025-01-13). <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan>.

型、AI框架到加速芯片的驱动没有形成自主支撑链，开源生态的主导权不在手上。美国参议员乔什·霍利2025年1月29日向美国国会提出《2025年美国人工智能能力与中国脱钩法案》(Decoupling U.S. Artificial Intelligence Capabilities from China Act of 2025)；若该法案得以通过，将彻底切断美国与中国在AI领域的合作。

(3) **头部企业的集群式号召力薄弱**。在应用创新领域，国内AI头部企业的技术优势和影响力尚不具备带动业内中小企业协同发展的能力，软硬件项目之间缺乏统一的兼容性标准和接口，技术“孤岛”现象突出，制约了生态的协作推进。和头部企业相比，一些新兴企业通过发布备受关注的开源产品和技术（如DeepSeek等），在社区中产生了重要影响，并表现了更强的创新能力和生态建设能力，已具备一定的引领号召能力，并建立了国产大模型的事实标准。

1.3 生态活力欠佳

(1) **开源人才供给面临短缺**。当前，我国对开源领域人才工作重视不够。受限于考核机制等影响，开源领域人才的培养未能得到足够的关注和支持，导致人才结构不够完善。具体而言，开源生态中缺乏从“关键运维者”到“核心贡献者”再到“一般贡献者”的完整人才梯队。这种结构性缺失使得我国开源生态难以持续获得高质量的专业人才支持，制约了开源创新生态的进一步发展。

(2) **生态对外扩展乏力**。国内的AI开源社区和开源代码托管平台主要以本土企业和研发机构推动，但缺乏具备全球推广潜力的基础产品，国际影响力和认可度较低，难以有效汇聚全球智慧。同时，政治因素也使国际环境更加复杂，进一步阻碍了全球合作。例如，在GitHub平台上，中国的开发者人数近年来增长显著放缓，并在2022年第1季度被印度超过，位居第

3位。2024年第3季度，中国和印度的GitHub开发者数量分别是996万和1 711万，相差近1倍^③。

(3) **高质量数据集严重匮乏**。不同数据集特征对模型性能有极大影响。随着AI大模型训练数据需求量的快速增加，高质量数据集逐渐成为稀缺资源。为了避免各种纠纷争议，国内外公开发布的大模型基本不附带相应训练数据集，出现模型算法开源与数据集专有闭源“倒挂”现象。国际上，知名的大语言模型训练数据集包括以Common Crawl为代表的通用领域数据集，以及PubMed和ArxivPapers为代表的专业领域数据集。在国内，我国虽然建设形成了各类数据中心，但是仍然缺乏专门面向大语言模型训练的高质量语料库和数据集，严重制约我国AI的发展。

1.4 生态运行机制不成熟

(1) **生态分工协作机制尚不完善**。国内AI开源合作多集中于“高校院所—企业”和“企业—开源组织”，“高校院所—企业—开源组织”合作链条尚未打通，难以形成合力。开源社区与专业化服务机构缺乏必要协作导致专业化、机构化运营治理程度不高，跨平台、跨项目协作机制尚不完善。缺少策源性AI开源组织和开源项目，导致我国从“0到1”的原始创新比较乏力。

(2) **AI开源的商业闭环尚未畅通**。尽管开源AI在技术上取得了显著进展，但商业化的成功案例相对较少。大多数开源项目侧重于社区建设和技术共享，而不是商业盈利。许多项目依赖捐赠、政府资助或企业赞助来维持运营，即使想商业化还面临知识产权保护、技术支持和市场推广等方面的挑战。开源大模型缺乏可持续的盈利途径。

(3) **在国际开源组织中话语权不足**。近年来，尽管国内AI领域的企业积极寻求与国际开源基金会等组织的合作，但往往停留在较浅层次，合作深度有限，

^③ Github. Developers. [2025-03-10]. <https://innovationgraph.github.com/global-metrics/developers>.

且在国际专业会议中的参与度较低。同时，政府、企业、科研院所和公益组织等多方主体尚未充分发挥各自优势，未能形成协同参与国际开源事务的多元化格局，因此限制了我国在全球开源生态中的整体竞争力。缺乏像欧盟 AI Watch 和开源观测站（OSOR）之类的长期跟踪国际 AI 和开源政策方面的情报平台，难以为国家战略决策提供决策支撑。

2 加快构建我国 AI 开源创新生态体系的建议

2.1 加强顶层设计，构建统合度高、协同性强的政策体系

（1）完善政策体系。制定 AI 开源创新生态建设的顶层规划和支持政策，明确发展目标、重点任务和保障措施，形成“顶层设计—专项政策—具体措施”系统性政策体系，积极融入国家层面的 AI、新型信息基础设施和开放科学行动方案。建立健全开源生态激励和利益分配机制，对创新主体的开源生态建设贡献进行全面评价，并在评价基础上采取多元化的激励方式，激发生态活力。

（2）强化政策协同。统筹协调各级政府部门，制定具体规范和实施细则，明确政策执行主体、责任分工和操作流程，加强政策衔接和配套，形成政策合力，避免政策碎片化和重复交叉，确保政策落地见效。在技术发展的原生阶段，政府应通过政策引导为市场创造良好的环境，尊重市场规律，充分发挥市场“无形之手”的力量，调动社会资本和群体智慧的积极性。在监管方面，政府应采取适度宽松的策略，以鼓励创新为主要导向，减少过度干预，从而促进开源技术生态的健康发展，推动技术创新和产业繁荣。

2.2 加快开源开放 AI 基础设施建设，夯实创新生态发展底层支撑

（1）构建开放协同的 AI 公共基础设施平台。联合政府、企业、科研机构 and 公益组织等力量，共同建设开源代码托管平台、开源大模型平台、开源数据平台等，

为开源项目提供开发、测试、训练、部署全流程支持。推动平台资源的互联互通、方便获取、易操作性和实惠价格，协同推和融入国家“新型信息基础设施”的建设和发展。

（2）加强开源硬件生态建设。重点发展高性能计算芯片和 AI 芯片等自主可控的芯片生态，以及支持高速计算处理及快速数据流通等硬件设施，为开源大模型提供强大的硬件基础。推动算力网络和算力调度技术发展，提升算力资源利用效率，满足 AI 应用需求。

（3）推动开源软件生态发展。支持开源操作系统、开源数据库、开源大模型、开源开发工具等软件的研发和应用，构建完善的软件生态系统，降低 AI 项目开发门槛；加强开源相关方（包括产业界、科研界、教育界和社会组织等）合作伙伴关系的发现、建设和扩展。以科研界为例，国家科学数据中心、国家资源库、重大科研基础设施和大型科研仪器等科技基础设施中就包含大量与开源相关的工作。支持新型研发机构或基金会组织构建完整的 AI 软硬件技术栈和工具集。

（4）加强 AI 开源基础设施在科研、教育和行业领域的应用推广。截至 2024 年 3 月，我国已批准 23 个国家新一代人工智能开放创新平台，这些平台在推动 AI 技术创新和产业应用方面发挥了重要作用。然而，面对当前快速演进的大模型技术生态，我国仍然缺乏一个面向全球开源开放协作、具有专业性和中立性的重大科技基础设施。该基础设施应能够整合并服务相关产学研单位，促进技术成果的共享与转化，推动多样化的应用场景示范工作，从而全面提升我国 AI 的科技基础能力水平。

2.3 培育多元参与主体，激发开源生态体系活力

（1）优化人才培养和激励机制。据行业报告的估计，到 2030 年，中国 AI 人才缺口预计将达到 400 万人。优化人才培养和激励机制，大力推广开源文化，加强人才政策的制定和实施。一方面，要加强本土人

才的发现、培养和成长；另一方面，要加大对全球人才的吸引。从OpenAI和xAI公司技术团队中频频亮相的华人面孔可以看出，华人在全球AI领域中的重要贡献和地位，我国应加强对高级AI人才的激励和引入，充分发挥他们在国内AI发展中的作用。

(2) 支持新型研发机构发展。鼓励企业积极参与开源项目，贡献代码和经验，并通过开源社区获取技术和人才支持，提升自身竞争力。加大对新型研发机构的支持力度，发挥其在AI领域智力资源优势，推动科研成果转化和开源生态建设。

(3) 加强数据集开源开放以及与数据集责任方的合作。国际数据公司（IDC）发布“数据时代2025”（Data Age 2025）的报告显示，到2025年，中国数据总量预计将跃居世界第1位，全球占比有望达到27%以上。但是，数据的开放共享和交互流通仍然存在诸多问题。制定数据开放共享政策，明确数据开放范围、标准和流程，鼓励政府、企业和科研机构合作，协同开放和维护高质量数据集，建设开源数据平台，促进数据资源共享和协同创新，有效应对高质量数据集短缺问题。积极响应国家《“数据要素×”三年行动计划（2024—2026年）》，积极建设国家大模型语料库，促进新质生产力快速发展。

2.4 完善开源创新运行机制，促进生态系统健康发展

(1) 建立开源协同合作机制。打通“高校院所—

企业—开源组织”合作链条，促进产学研深度融合。加强开源社区与专业化服务机构协作，提升运营治理能力。完善跨平台、跨项目协作机制，促进国内外资源共享和协同创新。

(2) 完善科技成果转化机制。推动基础研究与工程实践的紧密结合，加快开源和数据领域在知识产权及成果转化方面的制度建设。通过分离知识产权与使用权、数据集与模型算法，促进各方资源的互补与合作，打造“有限共享、无限合作”的创新生态。建议以DeepSeek为核心和契机，发起一个专注于下一代AI基础设施的基金会组织，旨在协调相关成果快速转化，并持续推动开源创新生态发展。

(3) 建立健全开源治理机制。打造AI开源开放一体化平台，建立健全开源生态协作与治理机制，加强在数据安全、数据隐私、算法偏见、法律法规、伦理责任等方面的合作应对；携手推动和落实中国2023年发起的《全球人工智能治理倡议》，以及2025年2月由中国和法国等61个国家共同签署发布的《关于发展包容、可持续的人工智能造福人类与地球的声明》。

(4) 优化国际创新合作机制。加强“破圈”行动，加强与开源模型、开放数据、开放文献、开放教育等相关工作的合作及应用案例培育和推广。积极参与和支撑开放科学、数字公共产品和AI造福人类等密切相关的国际行动计划，为联合国可持续发展目标等全球人类共同目标贡献优秀案例和中国方案。

Suggestions on building China's artificial intelligence open source innovation ecosystem

LONG Yuntao^{1,2} LIU Haibo^{1,2} XU Zheping^{3,4,5} BAO Yungang^{6*} WU Yanjun⁷

(1 Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China;

2 School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100190, China;

3 National Science Library, Chinese Academy of Sciences, Beijing 100190, China;

4 Key Laboratory of New Publishing and Knowledge Services for Scholarly Journals, Beijing 100190, China;

5 Department of Information Resources Management, School of Economics and Management,
University of Chinese Academy of Sciences, Beijing 100190, China;

6 Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;

7 Institute of Software, Chinese Academy of Sciences, Beijing 100190, China)

Abstract With the release and application of open-source large language models (LLMs) represented by DeepSeek, open-source innovation has played an increasingly important role in the rapid development of global artificial intelligence (AI) technology. Through technological innovation, cost reduction, performance improvement, and application demonstration, DeepSeek has opened up a disruptive innovation path and improved the research and application level of AI in China and its international influence. In order to further promote the construction of China's AI open source innovation ecosystem and promote the continuous innovation and iteration of AI technology, this study analyzes the outstanding problems in China's AI innovation ecosystem and the deficiencies of relevant policies, and puts forward suggestions to accelerate the construction of China's AI open source innovation ecosystem from four dimensions: top-level design, infrastructure construction, cultivation of multiple participants, and improvement of open source innovation operating mechanism.

Keywords artificial intelligence, open-source innovation, infrastructure, ecological system construction

隆云滔 中国科学院科技战略咨询研究院副研究员。中国计算机学会开源发展委员会执行委员、开源战略工作组组长。主要研究方向为开源创新发展、数字经济。E-mail: yuntaol@casisd.cn

LONG Yuntao Associate Researcher of the Institutes of Science and Development, Chinese Academy of Sciences (CAS), Executive Member of the Open Source Development Committee, China Computer Federation, and Head of the Open Source Strategy Working Group. Her research focuses on open source innovation development and digital economy. E-mail: yuntaol@casisd.cn

包云岗 中国科学院计算技术研究所副所长、研究员,中国科学院大学计算机科学与技术学院副院长、教授。《中国科学院院刊》青年编辑委员会主任,中国计算机学会开源发展委员会副主任,RISC-V国际基金会理事。主要研究方向为计算机系统结构。E-mail: baoyg@ict.ac.cn

BAO Yungang Professor, Deputy Director of the Institute of Computing Technology, Chinese Academy of Sciences (CAS), Deputy Dean of School of Computer Science and Technology, University of Chinese Academy of Sciences (UCAS), and Director of the Youth Editorial Board of *Bulletin of Chinese Academy of Sciences*, Vice Chairman of the Open Source Development Committee, China Computer Federation, Board of Director of RISC-V International. His research focuses on computer architecture. E-mail: baoyg@ict.ac.cn

■ 责任编辑：岳凌生

*Corresponding author