GOSD

2026

全球开源发展报告

GLOBAL OPEN SOURCE
DEVELOPMENT REPORT

从普惠到质量，从参与到影响。

持续共同创造及其公共价值



CCOSI
中国科协联合国咨商开源创新专业委员会
CAST UN Consultative Committee on Open Source Innovation

INCLUSION · QUALITY
PARTICIPATION · IMPACT

阶段合订版 · 前置内容与第一章 | 2026.10.22

全书目录

从共同愿景进入研究，以四个视角观察生态，再把证据带回行动。目录列出六章、三十个小节及两篇附录。

执行摘要 III

引言：开源共创智能未来 V

第1章 从开放协作到公共价值 1

1.1 开源：共同创新与公共价值 2

1.2 OSDGs：面向智能时代的开源发展目标 7

1.3 OSDI：指标、证据与观察方法 10

第2章 全球开源生态观测地图（GOSO）

2.1 全球开源活跃地图

2.2 全球开源贡献地图

2.3 全球开源项目影响力地图

2.4 全球开源人才地图

2.5 全球开源资产地图

2.6 全球开源多样性地图

2.7 全球开源治理与社会公益地图

第3章 全球开源影响力排行榜（GOSR）

3.1 全球行政区划影响力排行榜

3.2 全球开源企业影响力排行榜

3.3 全球开源基金会影响力排行榜

3.4 全球开源项目影响力排行榜

3.5 全球新势力开源项目排行榜

3.6 全球开源大模型影响力排行榜

全书目录（续）

第4章 全球开源技术全景洞察（GOSL）

4.1 Agentic AI 全景洞察

4.2 AI 大模型全景洞察

4.3 云基础设施全景洞察

4.4 数据库系统全景洞察

4.5 RISC-V 全景洞察

4.6 PyTorch 基金会全景洞察

第5章 全球开源大模型发展洞察（GOSM）

5.1 供给与贡献

5.2 网络与影响力

5.3 开源开放与分级体系

第6章 未来展望与行动建议

6.1 核心洞察与趋势展望

6.2 面向政府与国际组织的建议

6.3 面向企业与组织的建议

6.4 面向社区与全球合作的倡议

6.5 结语：迈向可持续的数字未来

附录与方法

附录 A：指标、数据与研究方法

附录 B：模型开放条件与数字公共品研究参照

本下载收录封面、目录、执行摘要、引言与第一章。其余正文与两篇附录可在线阅读；独立篇章 PDF 及附录 A＋B 合订 PDF 见对应篇章与下载页。

本版第四章重点展开 4.1；4.2—4.6 提供技术全景与研究框架，后续分析待补。其余章节中的样本范围、数据状态和模拟内容按图说明。

先从摘要把握主要观察，从引言理解共同愿景，再进入正文；也可按问题直接查阅。

本报告由中国科协联合国咨商开源创新专委会牵头，连接开源实践与可持续发展议题。研究以四大支柱组织观察，用地图、榜单、技术全景和模型生态分析理解发展质量，推动持续共同创造及其公共价值。

执行摘要

从普惠到质量，从参与到影响

人工智能的发展，使开源的协作对象从代码进一步延伸到模型、数据和智能体。开放成果不断积累，新的参与者与技术方向持续出现。与此同时，供给能否被理解和复用、关键基础能否得到维护、更多人能否进入共同创新，成为判断开源发展质量的重要问题。

《2026 全球开源发展报告》的核心主张是：将扩大参与和开放供给，与持续贡献、共同维护及公共价值连接起来。报告以开源发展目标（OSDGs）和开源发展指标（OSDI）为共同框架，通过生态地图、影响力榜单、技术全景和模型生态分析，观察开放成果怎样成为下一次创造的基础。发展成效既要看规模，也要看协作能否延续、成果能否积累、受益机会能否扩大。

主要发现与判断

参与基础与协作影响，需要分别观察。

第二章地图与第三章榜单表明，参与者数量与协作网络中的影响力位次并不总是一致。较大的开发者基础提供了发展条件，将其转化为持续贡献和项目影响，还需要真实任务、人才成长与组织支持。评价地区发展，应同时观察参与、贡献、项目与核心人才，避免将人数或单一排名作为全部成效。

开放生态的持续影响，依赖组织化协作与共同维护。

区域、企业、基金会和项目榜单揭示了不同主体在协作网络中的位置。企业连接研发资源与产品实践，基金会与社区承接跨组织协作，维护者让共同成果持续可用。由此，影响力观察应进一步连接责任与资源：哪些基础被反复使用，谁承担维护，使用者如何回馈上游。榜单提供识别对象的入口，长期支持仍应结合具体维护需求。

AI 开源生态的建设重点，正在扩展到完整技术体系。

第四章的 Agentic AI 全景将项目放回智能体与模型基础设施的技术分类中，呈现任务入口、编排、工具、记忆、评估与运行基础之间的分工。第三章中 PyTorch、vLLM 等项目的生态位置，也提示基础技术的重要性。对公共部门与产业而言，支持方向应覆盖可复用的基础工具、互操作与公开评测，使应用创新拥有稳定的共同基础。

模型供给迅速扩张，复用关系成为理解生态的重要线索。

按 2026 年 9 月同一快照中仍可见模型的创建时间回溯，模型数量较上年同日增长 58.3%，派生模型增长 52.3%。在第五章另一个覆盖 826 个组织账号的地域样本中，中美合计占模型数量的 71.7%；可识别的跨国派生记录同时呈现基座供给与下游开发之间的连接。这些证据提示，模型生态需要同时观察供给规模、发布主体和后续创造，支持既要回应关键基座的维护需求，也要为更多地区与团队保留参与机会。上述快照与样本分别解释，不能视为全球完整存量或国家贡献总量。

开放许可、完整交付与公共价值，需要建立各自的证据。

第五章的 870 个评估样本中，600 个归入高开放组，同时达到高完整与高开放的只有 117 个，占全体样本的 13.4%。许可条件与技术材料并不同步，公开成果也不自动意味着能够复现或产生公共收益。政策与采购应面向具体版本和用途，分别核验材料、许可、运行验证及真实使用结果，把开放承诺落实为可持续的协作条件。

三项共同任务，推动发展质量提升

基于这些观察，报告提出三项相互支撑的共同任务，并在第六章分别落实为政府与国际组织、企业与组织、社区与全球合作的九条建议。

支持持续参与与人才成长。

把人才支持连接到真实共建任务，为新人提供导师、文档和远程协作机会，为持续贡献者提供时间与认可，为维护者提供接续支持。公共部门改善参与条件，企业将贡献纳入研发安排，社区建立从首次贡献到共同维护的成长路径。成效应体现在持续贡献与责任成长，而不止于注册或活动人数。

建设可维护、可复用的开放技术基础。

把支持从首次发布延伸到文档、测试、兼容适配、问题响应和长期维护。企业应识别关键上游并分担维护责任，公共资助与基金会可支持跨项目共用的工具、评测和复现条件，社区组织上下游共同工作。资源配置兼顾关键基础与服务特定语言、地区和任务的长尾项目。

推进透明治理与可验证的公共价值。

明确版本、材料、许可和维护责任，开放重要协作规则与决策记录。将教育、无障碍和公共服务等需求转化为由使用者共同参与的具体任务，以真实采用、受益范围与后续维护检验成果。国际合作可围绕共同任务、共享证据和双向反馈开展，让本地经验进入上游改进。

本报告最终关注的是持续共同创造及其公共价值：让更多人能够参与，让成果在维护与复用中不断发展，让开放协作的收益能够被解释、验证和延续。

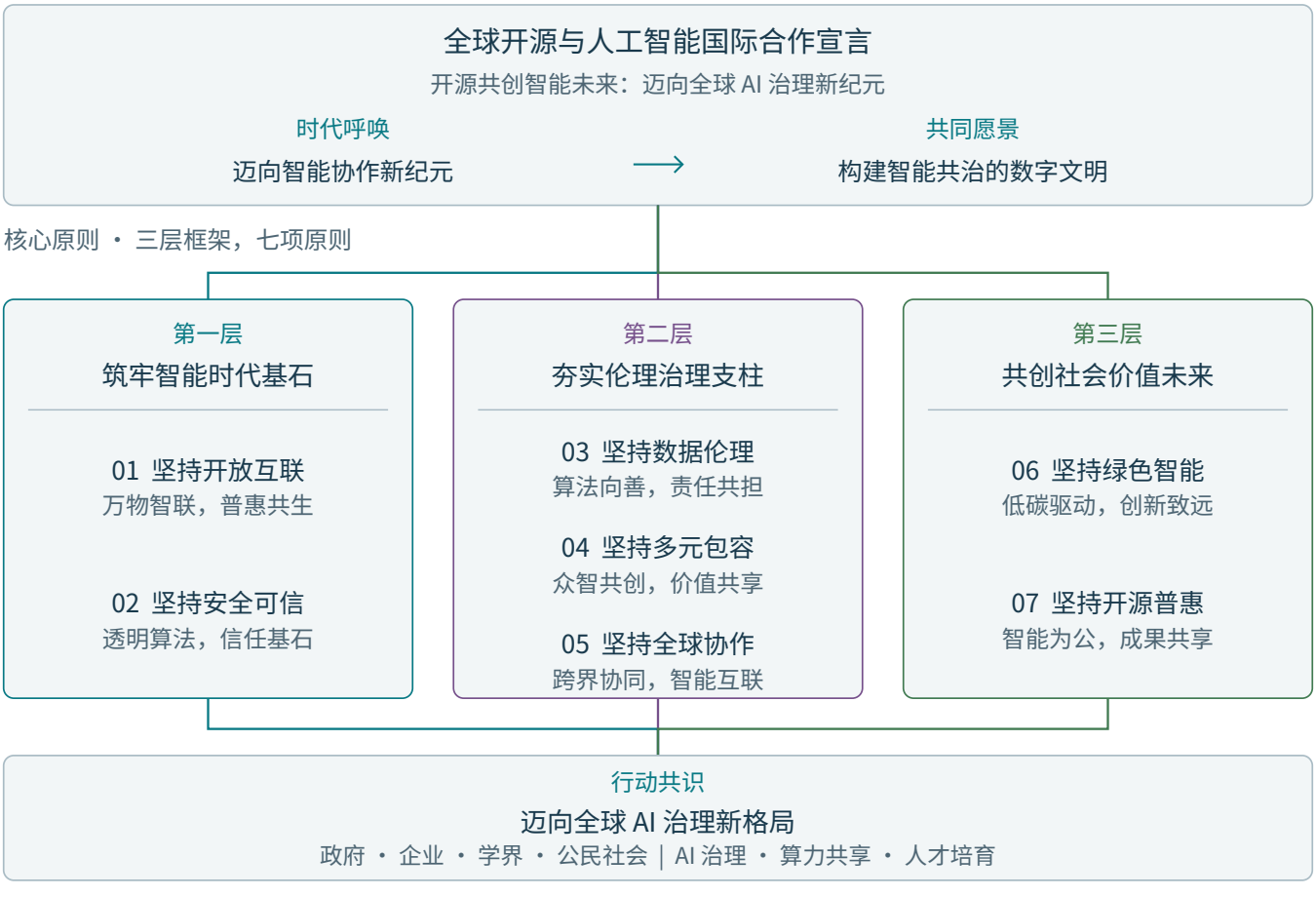
研究以 2025.07—2026.06 为主统计期；模型快照等证据按各节标注日期独立解释。模拟内容不作为事实结论的依据，未完成核验的指标不视为已测算结果。

引言：开源共创智能未来

本报告以《全球开源与人工智能国际合作宣言》为思想起点。宣言围绕开放互联、安全可信、数据伦理、多元包容、全球协作、绿色智能和开源普惠七项原则，提出建设开放、包容、可持续的人工智能生态的共同愿景。

《2026 全球开源发展报告》把这一愿景连接到证据与行动：谁能够参与，开放成果如何得到维护与复用，协作怎样形成可以验证的公共价值。下文宣言提供价值起点，后续章节进一步展开目标、观察与行动责任。

宣言表达共同愿景与行动原则；本报告以具体证据展开研究判断。二者相互衔接，价值主张与研究结论分别呈现。



开源创新，属于全人类；智能未来，由我们共同创造。

《全球开源与人工智能国际合作宣言》

—— 开源共创智能未来：迈向全球 AI 治理新纪元

I. 时代呼唤：迈向智能协作新纪元

我们正身处人类文明从“计算互联”向“智能共生”跃迁的关键节点。人工智能技术的爆发式增长，特别是大模型与 AI 智能体的崛起，以前所未有的深度重塑着全球经济与社会结构。然而，全球智能鸿沟持续扩大，技术主权博弈加剧，数据孤岛与算法黑盒成为悬在全人类头顶的达摩克利斯之剑。面对气候变化、公共卫生挑战以及复杂的国际关系，唯有坚持开源共享、协作共治，才能确保 AI 向善，避免技术垄断带来的不平等。

在此，我们 —— 来自中国政府、世界顶尖科技企业、国际学术机构、开源基金会及全球开发者社区的代表——郑重发布本宣言。我们呼吁以开放精神为指引，依托世界人工智能合作组织（WAICO）等多边平台，共建共享一个更加公平、透明、可持续的智能未来。

本宣言提出全球 AI 开源行动框架：以数字基础设施为基石，以实质贡献与伦理治理为支柱，以社会价值实现为使命，推动开源软件成为驱动智能时代可持续发展的核心引擎。

II. 共同愿景：构建智能共治的数字文明

我们致力于构建一个开放、包容、安全、可持续的全球 AI 开源生态，使强大的智能技术真正服务于全人类福祉。

我们坚信：开源不仅是代码的共享，更是信任的建立。在生成式 AI 与大模型时代，开源构成了全球数字公共品的底层架构。唯有坚持多元参与、透明可解释、贡献导向的协作原则，才能让技术创新成为推动和平、繁荣与可持续发展的共同力量。

我们共同倡导：将开源创新纳入全球数字公共品战略，作为实现《2030 年可持续发展议程》的关键路径，助力构建共治、共享、共赢的数字文明。

III. 核心原则：行动框架

第一层：筑牢智能时代基石

原则一：坚持开放互联

“万物智联，普惠共生”

打破算法壁垒与语言障碍，推动多模态模型的跨文化交互。鼓励支持多语言的本地化大模型开发，坚持技术发展的非歧视性。让发展中国家能够平等接入全球 AI 资源，使技术连接成为缩小数字鸿沟的桥梁，而非加剧不平等的工具。

原则二：坚持安全可信

“透明算法，信任基石”

反对互联网碎片化与技术封闭。支持开源作为验证算法安全、检测数据偏见的首选试验田。建立可信、弹性、可持续的 AI 供应链安全机制，确保关键基础设施的中立性与稳定性，为全球智能社会的运行筑牢底线。

第二层：夯实伦理治理支柱

原则三：坚持数据伦理

“算法向善，责任共担”

坚持数据隐私保护与算法公平性，防范技术滥用。推动AI治理遵循包容、敏捷、负责任的原则。鼓励开源社区探索可解释性 AI（XAI）的最佳实践，确保每一项 AI 决策都可追溯、可审计，始终服务于人类尊严与权利。

原则四：坚持多元包容

“众智共创，价值共享”

尊重文化多样性与认知能力的差异，反对任何形式的“算法歧视”。特别关注弱势群体在 AI 时代的权益，通过开源项目赋能边缘群体。支持多元化的开发者社群，让每一个声音都能在智能技术的演进中被听见、被尊重。

原则五：坚持全球协作

“跨界协同，智能互联”

视开源为连接全球智慧的网络。支持跨国界、跨机构的联合研发，特别是在应对气候变化、生物医药等全球性挑战领域。促进知识、算力与高质量数据集的自由流动，强化南南合作与多边主义，构建去中心化但互联互通的创新生态。

第三层：共创社会价值未来

原则六：坚持绿色智能

“低碳驱动，创新致远”

优先支持应对气候变化、环境保护、清洁能源等议题的开源 AI 项目。倡导高效的模型训练范式与算力调度优化，降低 AI 运行的碳足迹。推动开源技术作为数字公共品的核心实现路径，在能源管理、环境监测中的广泛应用，确保智能技术的发展不以牺牲生态环境为代价。

原则七：坚持开源普惠

“智能为公，成果共享”

推动开源技术广泛应用于政府治理、公共卫生、教育公平等领域，服务社会治理现代化与民生改善。鼓励开源成为数字公共品的载体，建立以贡献为导向的激励机制，让创新成果惠及所有人，尤其是边缘群体。

IV. 行动共识：迈向全球AI治理新格局

我们呼吁：各国政府、科技企业、学界与公民社会携手同行，以本宣言为基础，深化在 AI 治理、算力共享、人才培育等领域的务实合作，争取早期收获。

我们支持依托联合国及相关国际机制，围绕人工智能伦理准则、跨境数据流动规则、AI 版权保护等紧迫议题展开对话，凝聚共识，推动开源成为构建公正合理全球人工智能治理体系的生动实践。

结语

开源，是我们对开放精神的坚守，是对人类团结的信念，是对更美好世界的共同追求。让我们以协作代替割裂，以透明回应质疑，以共享消弭垄断。共同构建一个更加开放、公平、可持续的智能文明。

开源创新，属于全人类；智能未来，由我们共同创造。

从共同愿景，到证据与行动

本报告以“从普惠到质量，从参与到影响”回应宣言提出的方向：以 OSDGs 明确发展目标，以 OSDI 组织观察；通过生态地图、影响力榜单、技术全景和模型生态理解共同创造，再由第六章提出三项共同任务与九条主体建议。

宣言表达共同追求，报告用证据观察进展与缺口。我们希望让更多人能够参与，让共同成果得到持续维护与复用，最终支持持续共同创造及其公共价值。

宣言依据报告收录文本；指定来源：飞书文档《全球开源与人工智能国际合作宣言》。英文为本报告译文。

第一章 从开放协作到公共价值

开源怎样让更多人参与技术创造，怎样把分散的贡献连接成可以持续发展的技术生态，又怎样产生可验证的公共价值？引言提出共同愿景，本章从具体案例出发，为这些问题建立共同的研究语言。

开放发布使知识和技术有机会被更多人使用，但成果能否延续，还取决于发布之后的参与、维护与再创造。代码需要审查和修复，模型与数据需要清晰的使用条件，新的参与者也需要学习、反馈和承担责任的机会。理解开源发展，既要看见成果的规模，也要看见支撑这些成果的人、组织和协作关系。

人工智能进一步拓展了这一观察范围。代码之外，模型、数据、开发工具和智能体成为相互联系的开放资产；协作既发生在一个项目内部，也沿着采用、修改和后续发布向外延伸。成果数量、网络影响和开放程度分别回答不同问题，任何一个数值都不足以独立说明整个生态的发展质量。

本章依次回答三个问题。1.1 通过 Linux、开放模型与数字公共品，讨论共同创新如何连接公共价值；1.2 以开源发展目标（OSDGs）的四大支柱、十个维度说明希望推动的发展方向；1.3 以开源发展目标指标体系（OSDI）的二十九项指标说明如何组织证据、比较差异与辨认局限。

宣言的七项原则表达共同价值，OSDGs 组织发展目标，OSDI 定义具体观察。三者相互衔接，不作逐项对应，也不合成为一个总分。借鉴联合国可持续发展目标的思路，是为了使问题与证据相联系，不意味着联合国对本报告框架、指标或排名的认证。全书研究路径见图 1-0-1。



图 1-0-1 全书结构与研究路径

全书由此形成“建立基础、展开观察、连接行动”的阅读路径。第二章观察全球开源生态的规模、分布与联系，第三章辨认形成影响的地区、组织和项目，第四章呈现技术分类与演进，第五章追踪模型的供给、复用与开放。四个研究视角相互补充，第六章再将证据汇聚为共同任务与主体建议，落到持续共同创造及其公共价值。

读者可以顺序阅读，也可以按问题进入相应章节。无论从哪幅图开始，都需要先确认观察对象、统计时间和指标含义，再判断它能够支持什么结论。第一章提供的正是这一共同起点。

1.1 开源：共同创新与公共价值

1.1.1 从一次发布到持续创造

一项技术成果公开发布，是开放协作的起点。此后仍需要有人回应问题、审查修改、修复缺陷和维护文档，也需要让新的参与者逐步理解项目并承担工作。成果能否被持续使用，取决于这些日常贡献能否延续。Linux 为理解这种共同维护提供了一个具体入口。

报告采集的 LFX Linux Kernel 累计快照记录了 45,855 位贡献者与 5,686 个组织条目，采集时点为 2026 年 10 月 11 日。图 1-1-1 分别展示组织与贡献者的前十条记录。组织榜单中，Red Hat 的贡献记录量为 725,866，占该榜单统计总量的 14.95%；Intel Corporation 为 537,913，占 11.08%。贡献者榜单中，David S. Miller 的记录量为 236,673，占该榜单统计总量的 3.73%。

5,686 组织条目
累计榜单总记录数

45,855 贡献者记录
累计榜单总记录数

A 组织榜单前 10 条

		贡献记录量	本类榜单占比
01	Red Hat	725,866	14.95%
02	Intel Corporation	537,913	11.08%
03	Linux Kernel	277,346	5.71%
04	Linaro	233,646	4.81%
05	SUSE	200,253	4.12%
06	Google	193,149	3.98%
07	Advanced Micro Devices, Inc.	161,146	3.32%
08	IBM	133,660	2.75%
09	Meta Platforms, Inc.	109,011	2.25%
10	Oracle America Inc	108,617	2.24%

B 贡献者榜单前 10 条

		贡献记录量	本类榜单占比
01	David S. Miller	236,673	3.73%
02	Linus Torvalds	193,264	3.05%
03	Greg KH	164,535	2.59%
04	Mark Brown	113,560	1.79%
05	Alex Deucher	93,428	1.47%
06	Mauro Chehab	90,300	1.42%
07	akpm	68,709	1.08%
08	Jakub Kicinski	67,230	1.06%
09	Ingo Molnar	65,226	1.03%
10	Takashi Iwai	47,625	0.75%

LFX Linux Kernel 累计快照 · 2026-10-11。组织与个人占比分别计算；条形按各自榜首缩放。

图 1-1-1 Linux 组织与贡献者累计排行

这两类榜单提供了互补的观察尺度：组织记录使资源投入与机构参与更可见，个人记录则把具体贡献者带入视野。来源中的“组织”是平台归集的条目，不能直接等同于企业；“贡献记录量”也应按来源定义理解，不能改称代码行数、OpenRank 或经济价值。组织与个人采用不同的统计分母，二者的占比不能相加，榜单记录也不能直接互相换算。

共同维护，让成果能够延续。累计榜单首先告诉我们“谁留下了贡献记录”，并为研究协作结构提供入口。它不能独立证明每位参与者在当年仍然活跃，也无法仅凭一个时点判断维护责任是否稳定。要回答这些问题，还需结合跨期活动、审查与发布记录，以及维护者分工等材料。

这一案例提示我们，支持开源不能只围绕新项目和新版本展开。文档改进、问题响应、版本维护和参与者成长，同样需要长期投入。第二章对参与和协作格局的观察、第三章对贡献主体的比较，将沿着这一思路展开；规模与排行是认识生态的起点，持续参与和责任延续是需要继续追踪的问题。

1.1.2 从代码协作到模型与数据的再创造

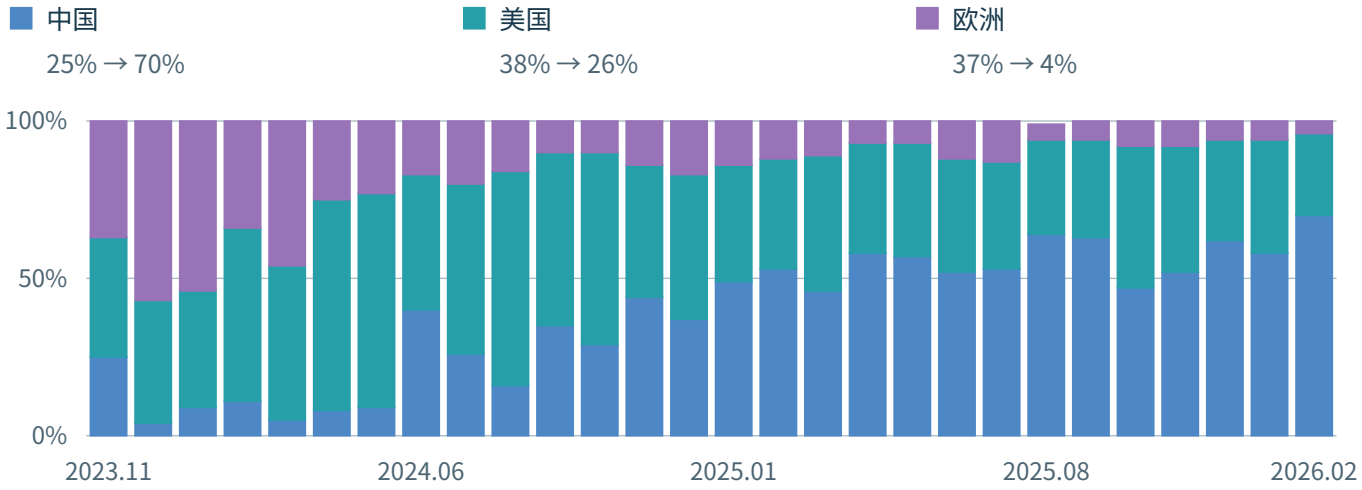
开放的影响也会沿着后续开发延伸。在模型生态中，基础模型可以成为指令微调、领域适配和其他开发活动的起点，后续成果又可能支持新的使用场景。共同改进一个成果，以及基于已有成果继续创造，构成两种相互补充的协作方式。后一种关系提醒我们：一项成果的影响，可能超出它自身的发布页面。

图 1-1-2 的宏观部分展示 ATOM 跟踪的 Hugging Face 样本。2023 年 11 月至 2026 年 2 月，新增微调与衍生模型所采用的基础模型，其发布方地区构成持续变化。样本起点，中国发布方占 25%、美国占 38%、欧洲占 37%；到 2026 年 2 月，三者分别为 70%、26% 和 4%。中间月份存在多次波动，期末差异不能被读成一条持续、匀速增长的路径。

这里观察的是基础模型在新增衍生中的构成，归属指基础模型发布方，不是下游用户国籍。占比上升说明其在该样本当月构成中的位置变化，不能仅据此判断衍生总量增长了多少、模型性能孰优，或整个市场由谁主导。原始百分比按来源保留，个别月份分项之和因取整并非恰好 100%。

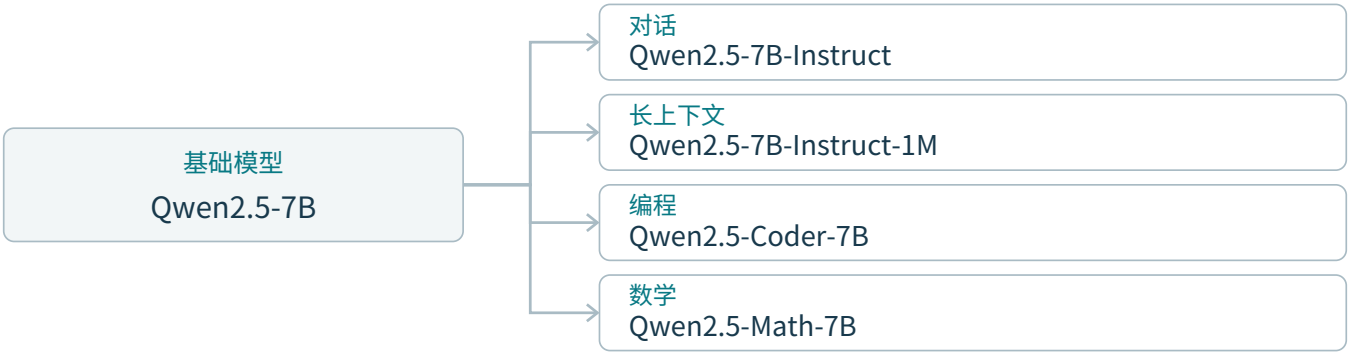
局部实例让这种复用关系更加具体。第五章保留的 2026 年 10 月 16 日平台关系快照中，Qwen2.5-7B 与 Qwen2.5-7B-Instruct、Qwen2.5-7B-Instruct-1M、Qwen2.5-Coder-7B、Qwen2.5-Math-7B 之间存在基础模型声明关系，分别连接对话、长上下文、编程和数学等用途。图中选取这四条直接关系，说明一次发布如何与后续创造相联系；它既不是该模型的完整下游清单，也不用于解释宏观图中的地区份额。

A 新增微调与衍生模型的基础模型发布方地区构成



ATOM 跟踪的 Hugging Face 样本，2023.11-2026.02。地区按基础模型发布方归属；图例列示首月与末月占比，取整可能使合计不等于 100%。

B Qwen2.5-7B 与四个选定的直接下游



关系快照：2026-10-16。连线为发布者声明，展示四个选定下游。

图 1-1-2 开放模型的宏观趋势与 Qwen 衍生实例

可复用的成果，连接下一次创造。这些连线记录的是发布者声明，并不验证实际训练过程，也不分配各方贡献比例。宏观样本与局部实例采用不同时间和范围：前者帮助观察构成变化，后者帮助理解关系含义，二者不能直接拼接成因果解释。

复用还依赖开放条件。能下载模型权重，不意味着训练数据、训练代码和评估材料都可获得；能运行模型，也不意味着使用、修改和分发没有条件。因此，模型数量、复用关系与开放程度需要分别观察。第五章以供给、网络与开放分级三个小节展开这些问题，将“发布了什么”进一步连接到“如何被继续使用与创造”。

1.1.3 SDGs 与数字公共品：连接发展目标与实践条件

教育、科研和公共服务可以采用开放成果，但技术可获得与实际受益之间，还存在使用条件、维护能力和本地适配等环节。对公共投入而言，值得追问的不只是发布了多少成果，还包括谁能够使用、谁承担维护，以及受益能否持续。公共价值需要沿着使用过程继续观察。

联合国可持续发展目标（SDGs）提供共同的发展方向。优质教育、产业与创新、减少不平等、有效治理和伙伴关系等议题，帮助我们提出“希望改善什么”的问题。将开源与这些目标联系起来，是为了明确研究和行动的方向；具体项目是否改善了某项发展结果，仍需相应场景中的证据。

数字公共品联盟（DPGA）的标准从另一侧回答“需要具备什么条件”。九项要求涵盖 SDG 相关性、开放许可、清晰权属、平台独立性、文档、非个人识别数据提取、隐私与适用法律、标准与最佳实践，以及不造成伤害。图 1-1-3 将十七项目标与九项条件并列呈现，作为理解发展方向和实践要求的两个入口，二者不作逐项对应。

SDGs · 17 项发展目标



01 无贫穷
02 零饥饿
03 良好健康与福祉
04 优质教育
05 性别平等
06 清洁饮水和卫生设施
07 经济适用的清洁能源
08 体面工作和经济增长
09 产业、创新和基础设施
10 减少不平等
11 可持续城市和社区
12 负责任消费和生产
13 气候行动
14 水下生物
15 陆地生物
16 和平、正义与强大机构
17 促进目标实现的伙伴关系

DPGA · 9 项实践条件

01 SDG 相关性 说明与可持续发展目标的关联。	02 开放许可 提供允许复用的开放许可。	03 清晰权属 说明数字成果的权属。
04 平台独立性 说明对封闭平台的依赖。	05 文档 提供理解和采用所需的资料。	06 非个人识别数据提取 在适用时支持非个人识别数据提取。
07 隐私与适用法律 披露隐私及适用法律要求。	08 标准与最佳实践 遵循相关标准和实践。	09 不造成伤害 数据安全、有害内容与骚扰防护。

DPG 条件据 DPGA Standard 简述，CC BY-SA 4.0。

图 1-1-3 联合国 SDGs × 数字公共品联盟 DPGA

发展目标、开放条件与实际成效，需要分别判断。与 SDGs 相关，不等于已获得数字公共品认可；通过标准评审，说明核心解决方案符合相应要求，仍不能替代对地方实施成效的验证。本报告借鉴目标与条件组织研究，不意味着联合国或 DPGA 对本报告框架、指标和排名的认证。

例如，开放的教育软件可以提供复用基础，但学习者能否受益，还取决于语言适配、设备可及、教师支持和持续维护。这是理解证据链的示例，不是本报告已经验证的项目成效。观察公共价值，需要把技术供给、采用条件和实际结果分别记录，再审慎解释它们之间的联系。

1.1.4 将持续共同创造作为观察重点

Linux 让贡献主体和维护问题进入视野，模型实例让下游复用关系更加具体，SDGs 与数字公共品则把这些技术活动连接到发展方向和实践条件。三者共同提示：开放的价值在参与、维护、使用与再创造中逐步形成，需要持续观察。

本报告据此提出三项共同任务：支持持续参与与人才成长，建设可维护、可复用的开放技术基础，推进透明治理与可验证的公共价值。它们是统筹后续研究与行动的方向，而非由三个案例直接推导的量化结论。下一节将这些方向展开为发展目标，第六章再结合各章证据，说明政府、企业与社区可以承担的行动。

1.2 OSDGs：面向智能时代的开源发展目标

1.2.1 用共同目标连接分散的观察

开源生态的参与者关注不同问题。公共部门关心发展机会与技术供给，企业关心关键技术和长期协作，研究机构关注知识传播，社区与基金会关注维护和治理。OSDGs 的作用，是为这些问题建立共同的价值坐标，使不同主体能够讨论同一项技术生态的不同方面。

本报告沿用四大支柱、十个维度的目标框架。四个支柱分别关注参与基础、开放供给、连接与复用、治理与公共价值。它们共同描述希望推动的发展方向；各个维度再将这一方向转化为可以持续观察的问题。图 1-2-1 展示完整结构：先找到关心的目标，再沿着对应维度寻找证据。



图 1-2-1 OSDGs 四大支柱与十个维度

1.2.2 人才规模与多样性

人才是协作的参与基础。第一支柱同时观察参与规模、聚集密度、新人进入和协作分布。规模有助于理解参与基础的大小，多样性则帮助我们追问：哪些地区、机构与群体能够进入协作，新增参与能否获得支持。

“规模与密度”关注活跃开发者有多少、相对于人口有多密集，以及新人怎样进入；“生态多样性”关注跨地区协作和学术机构参与。二者共同帮助区分参与基础的扩大与参与结构的变化。人数较多的地区未必具有较高的人才密度，新增账号也不能直接当作新晋贡献者。

第二章从全球和中国的空间分布切入，第三章通过地区与组织榜单补充观察。比较地区时，规模和密度分别回答不同问题；比较参与群体时，也应考虑地域、机构归属的覆盖程度。新人进入、长期留存和核心维护者成长需要各自的数据，不能用一个活跃人数概括。

1.2.3 技术产出与开放供给

第二支柱将代码、模型和数据放入共同的供给视野。代码仓库、模型仓库与数据集是不同资产，应保留各自的统计单位、发布方式和开放条件。供给规模帮助我们看见技术资源的积累，供给方多样性和成果结构则帮助我们识别选择空间与依赖关系。平台上公开可访问的仓库是供给观察的对象，其材料与许可是否满足相应开放要求，还需要另行核验。

三个维度分别是“代码产出”“模型开放供给”和“数据开放供给”。代码产出同时关注活跃仓库与协作贡献；模型和数据供给同时关注发布规模与供给方结构。把三类资产并列，可以辨认技术基础的不同环节；把它们简单相加，则会掩盖统计单位和复用方式的差别。

第四章的技术全景将项目放回技术体系，解释它们各自承担的功能；第五章从全球、中美和中国内部三个视角观察模型与数据供给。两者共同提醒我们，技术生态的完整性取决于不同环节的协同，也取决于关键成果是否具有持续维护和复用的条件。

1.2.4 全球连接与网络效应

第三支柱关注贡献和成果如何通过网络传播。代码项目之间的协作关系、模型的衍生关系、数据的采用与复用，分别展示知识如何进入更广泛的创新过程。影响力的观察由此需要保留网络结构，而不只统计孤立对象的数量。

“代码影响力”“模型影响力”和“数据集影响力”构成三个观察维度。代码侧关注协作网络及其中的项目、组织和人才；模型侧关注采用、社区参与和后续派生；数据侧关注采用和可核验的复用关系。一个成果被反复获取与成为后续创造的基础，是两类需要分别记录的信号。

第二、三章通过 OpenRank 等指标观察代码生态中的贡献与影响；第四章在具体技术分类中分析项目与类别的演进；第五章结合模型衍生网络和 ModelRank 月度资料，研究模型生态的连接。不同关系需要采用相应的度量方法，连接丰富程度也需要与集中性、依赖结构一起理解。

1.2.5 开源善治与公共价值

第四支柱关注协作能否持续，以及成果能否被负责任地使用。治理主体、维护规则和许可信息构成观察的重要入口；模型与数据进一步提出材料透明、使用条件和来源记录等要求。

“治理成熟度”关注谁承担治理和维护、参与是否集中、许可信息是否清楚；“公共福祉”关注数字公共品共建、开放条件与 AI 资产透明度。这些维度提供了可以追踪的入口，具体社会成效仍需要结合使用场景和受益者另行验证。

基金会托管情况、许可覆盖和模型开放分级可以提供不同方面的证据，但均不能独立代表治理的全部质量。公共福祉还需要识别具体使用者、服务场景与受益结果。本报告在证据充分的地方开展比较，在仍需长期研究的地方明确观察问题，保留公共价值的验证任务。

1.2.6 四个支柱如何共同工作

四个支柱之间存在联系，也保留各自的判断空间。参与扩大之后，可能仍然存在维护资源不足；供给增长之后，可能仍然缺少高质量文档和明确许可；一个影响力较大的项目，也可能面对贡献集中和持续维护的问题。框架的价值在于让这些问题同时可见。

第二至第五章共同使用四个支柱，不作一一配对。例如，模型生态同时涉及供给、多样性、网络效应和开放治理；一个技术全景也可以同时提出人才、项目影响与维护的问题。读者可以从章节进入具体领域，也可以从一个支柱跨章节寻找相关证据。

在行动层面，这些目标汇聚为三项共同任务：支持持续参与与人才成长，建设可维护、可复用的开放技术基础，推进透明治理与可验证的公共价值。不同主体可以据此找到自己的责任，并使用共同的观察框架检查进展。下一节通过 OSDI 说明这些进展如何被记录和比较。

1.3 OSDI：指标、证据与观察方法

1.3.1 用指标理解发展

OSDI 将四大支柱、十个维度展开为二十九项指标，从参与规模、开放供给、连接与复用、治理与公共价值等方面观察开源发展。图 1-3-1 将全部指标放回对应维度，说明各项观察在框架中的位置。

这些指标分为三类：北极星指标关注核心发展结果，驱动指标跟踪相关活动和投入，诊断指标帮助解释差异与约束。三类指标各有用途，不直接加总为一个总分。一个地区、组织或项目在不同维度上的表现，共同构成对它的理解。

北极星：核心结果　驱动：活动与投入　诊断：差异与约束

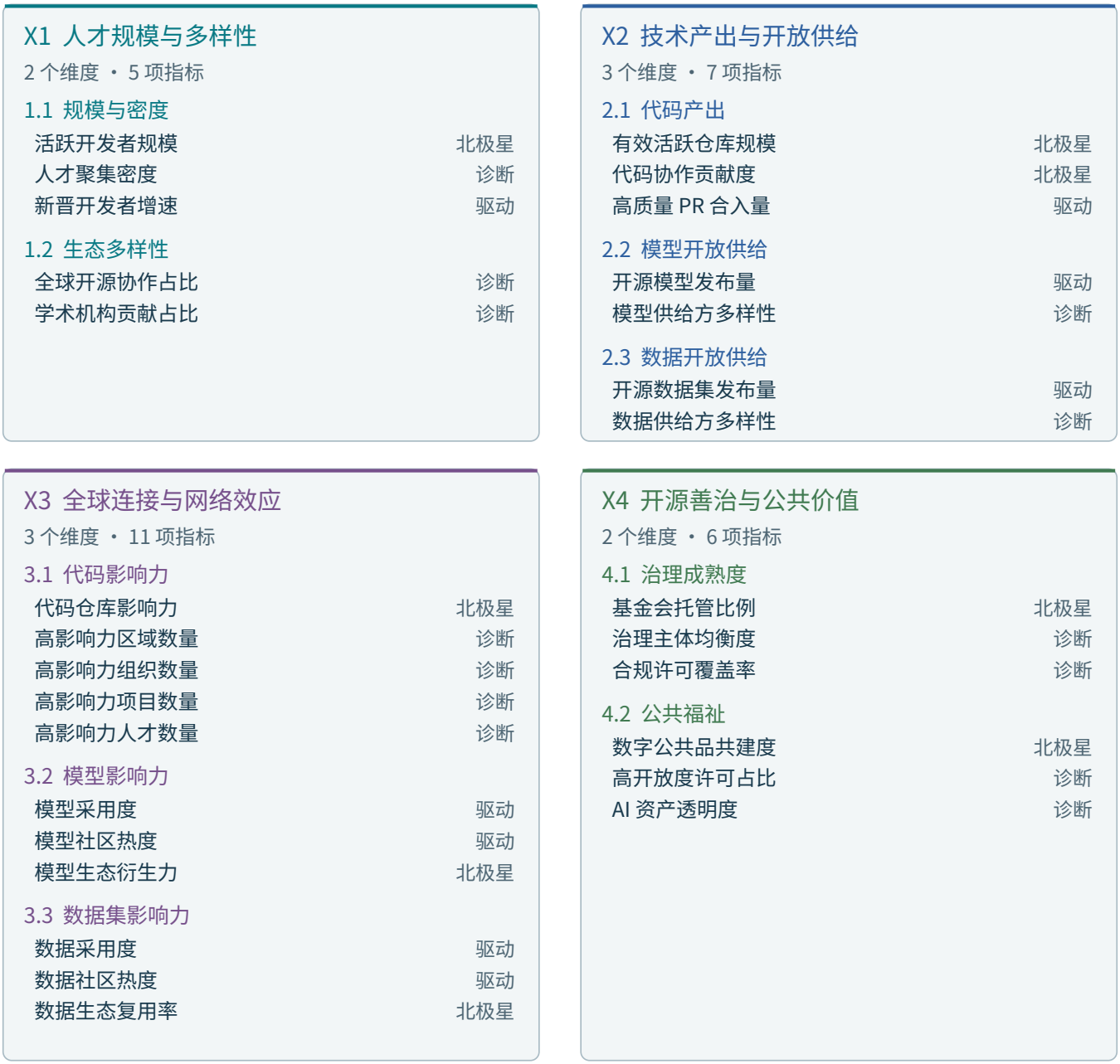


图 1-3-1 OSDI 指标全貌与三类用途

例如，在“人才规模与多样性”支柱下，活跃开发者规模帮助判断参与基础，新晋开发者增速跟踪进入变化，人才聚集密度帮助解释地区差异。在“模型影响力”维度下，ModelRank 观察模型作为后续开发基础的影响，模型采用度和社区热度补充解释相关活动。指标角色说明它被用来回答什么问题，不表示因果关系或计算权重。

二十九项指标是完整的观察框架，不意味着本期已完成全部测算。使用一项指标时，还需核对定义、统计对象、计算规则和数据状态。已有图表提供相关证据，只有在对象范围和计算方法一致时，才能进一步支持跨地区、跨年度比较。

1.3.2 从参与、贡献到影响与开放

读懂后续图表，可以先分清四个问题：谁在参与，贡献了什么，形成了怎样的影响，以及成果开放了什么、允许怎样使用。活跃人数、贡献量、影响力和开放等级分别回答这些问题，需要结合起来看。

以同一个项目为例，参与人数说明协作基础，代码提交、问题响应和评审记录说明具体活动，网络关系帮助观察这些活动怎样与其他参与者和成果相联系，材料与许可则说明后续使用的条件。增加一种观察，是为了补充认识；不能因为人数、下载量或影响力较高，就推定其他方面同样优秀。

代码生态的观察以 OpenDigger 的工具与指标为基础。OpenRank 通过协作网络刻画影响：社区 OpenRank 关注项目内部，全域 OpenRank 观察更广泛的开发者与仓库网络。按贡献者地域汇总，反映贡献来自哪里；按项目归属汇总，反映哪些主体的项目具有影响。两者各有统计对象，不能互相替代。

模型生态进一步关注供给、复用与开放。模型和数据集记录展示资产积累，基础模型声明关系展示后续创造，ModelRank 提供网络影响的观察。本期纳入了 2024 年 7 月至 2026 年 8 月共 26 个月的 ModelRank 资料；这些排名不代表模型能力测试，算法的独立复现仍需核验。

开放程度则单独评估。第五章依据 SJ/T 12304—2026《人工智能 关键基础技术 大模型技术体系开源开放分级评估规范》，分别介绍材料开放与许可开放。材料等级需满足本级及前置等级要求，许可条件另行判断。标准已发布，将于 2026 年 12 月 1 日实施；平台已有样本标签保留原评估版本，不视为本报告已按发布版标准重新评级。

1.3.3 在共同范围内比较，带着问题进入观察

比较之前，先确认对象、时间和范围。报告年度为上年 7 月 1 日至当年 6 月 30 日；自然年、月度与平台快照按图中标注阅读。当前样本按创建时间回溯的趋势，只反映样本内资产的积累，不能完整还原历史存量。

开发者、仓库、模型版本和发布组织是不同的统计单位。份额与排名只在明确的比较集合中成立，未知归属单列。下载量不等于实际用户数，仓库数不等于独立模型家族数，开放等级也不代表性能、安全性或社会效益。

指标有定义、数据已入库、图表已使用、计算可复现，是分别记录的状态。观测、派生、模拟与预测也逐图区分；模拟和预测用于展示与探索，不作为事实性排名或政策优先级的实证依据。治理与公共价值尤其需要结合具体场景、维护责任和受益结果来判断。

图 1-3-2 将四大支柱连接到第二章的七个主题地图：活跃、贡献、项目影响力、人才、资产、多样性、治理与社会公益。读者可以从关心的目标进入相关主题，也可以围绕同一问题交叉阅读。图中连线是阅读与研究联系，不是指标权重；一个主题可以回应多个目标。



连线表示阅读与研究联系，不代表计算权重。

图 1-3-2 从四大支柱到七个主题地图

带着这些问题，第二章从七个主题观察全球与中国的生态，第三章识别贡献与影响主体，第四章展开技术结构，第五章观察模型的供给、复用与开放，第六章汇聚共同任务与主体行动。全书最终关心的是：更多人能否持续参与，共同创造的成果能否得到维护和继续使用，其公共价值能否被验证与改进。