



2024 中国开源发展现状

2024 China Open Source Development Status

2024 中国开源发展现状

编写委员会

顾问: 陆首群

策划: 刘 澎 赵 琛 蒋 涛 梁志辉

主编: 孟迎霞 宋可为 隆云滔 王 伟 孙 启

荆 琦 谭中意 唐小引 鞠东颖

工作组 (按姓氏首字母排序):

安 泱 白国华 程智锋 邓 伟 丁禹民 韩凡宇

何 苗 梁华博 李 鑫 刘巍巍 贾朔维 宁 固

宋逸群 滕召智 屠 敏 王蕴博 许哲平 袁 新

章文嵩

目录

总论 ————— 5

 2024 中国开源发展的机遇与特色 ————— 6

 中国开源发展的风险和挑战 ————— 9

 中国开源发展建议 ————— 10

第一章 开源通识 ————— 12

 1.1 开源的定义和特征 ————— 12

 1.2 世界开源发展简史 ————— 13

 1.3 中国开源发展简史 ————— 18

 1.4 开源软件与知识产权 ————— 23

 1.5 开源治理 ————— 27

第二章 开源开发者及项目发展现状 ————— 32

 2.1 开源开发者发展现状 ————— 32

 2.2 开源项目发展现状 ————— 46

 2.3 开发者面临的挑战、建议与未来展望 ————— 55

第三章 开源重点领域发展现状 ————— 60

 3.1 操作系统产业研究报告 ————— 60

 3.2 数据库产业发展报告 ————— 67

 3.3 人工智能产业研究报告 ————— 81

 3.4 RISC-V行业研究报告 ————— 116

 3.5 智能网联汽车行业研究报告 ————— 123

 3.6 开源供应链研究报告 ————— 130

第四章 开源教育现状 ————— 143

 4.1 开源教育加速融入并贯穿教育全阶段 ————— 143

 4.2 开源教育平台和开源协作模式 ————— 144

 4.3 校企结合的开源教育协同模式 ————— 146

 4.4 开源教育问题与挑战 ————— 149

 4.5 开源教育发展建议 ————— 150

第五章 开源企业及机构发展现状 ————— 151

 5.1 中国开源企业发展现状 ————— 151

 5.2 开源基金会发展现状 ————— 173

 5.3 其他开源组织 ————— 176

第六章 开源国际化发展现状 ————— 179

 6.1 国际开源基金会和企业对中国的贡献 ————— 179

 6.2 中国企业对国际开源基金会的贡献 ————— 187

 6.3 COPU中沙开源合作进展及成果 ————— 195

附录1: 开源大事记 ————— 199

附录2: 2023-2024年度开源榜单汇编 ————— 218

附录3: 开源专有名词 ————— 222

附录4: 开放科学和开放标准 ————— 230

附录5: 开源合规案例精选 ————— 239

致谢 ————— 247

总论

在2024年的全球开源发展浪潮中，中国开源继续展现出强劲的发展势头和独特的创新活力。中国开源社区的活跃度持续提升，开发者和贡献者数量显著增长，开源项目在国内外的影响力不断扩大，企业参与度不断加深，以华为、阿里巴巴为代表的中国企业在国际开源项目中的贡献日益凸显。同时，国内开源基金会和组织建设，为开源项目的规范化管理和健康发展提供了有力支撑。中国已成为开源参与者数量全球排名第二、速度增长最快的国家。

回顾2023年，中国开源生态的各个方面，包括开源项目的发展、开源基金会的建设、知识产权的保护、商业模式创新、投融资活动的活跃以及公共服务平台的完善，都取得了快速的发展，共同推动了开源生态的繁荣和壮大。随着这些积极因素的相互作用和促进，中国开源不仅在国内形成了强大的创新动力，在国际开源舞台上也不断展现出其独特的价值和影响力。中国的开源实践正深度融入到国际开源生态之中，越来越多的中国企业已经成为了开源生态的市场主体，积极投身开源，并将更多的市场资源投入开源创新过程，全球影响力逐步扩大，部分主体已经成为引领全球开源生态建设的重要力量。

为更准确、客观、真实、完整地展现中国开源发展现状，由COPU牵头，联合中国开发者社区CSDN、中国科学院软件研究所、北京开源创新委员会、开源社、开源中国、北京大学、华东师范大学、国防科技大学等80多家单位以及102位开源专家和志愿者共同协作编撰完成这本《2024中国开源发展现状》，旨在全面展现中国开源产业生态的现状，深入分析中国开源在技术创新和产业发展方面的显著成就，并探讨当前所面临的机遇与挑战。

在2021-2023年度中国开源发展蓝皮书基础上，《2024中国开源发展现状》从6个章节分别介绍了开源通识、开源开发者及项目发展现状、开源重点领域发展现状、开源教育现状、开源贡献共生共赢等，尤其是对操作系统、数据库、人工智能、开源供应链、RISC-V、智能网联汽车等重点产业进行专项研究，最后以附录的形式对开源大事记、开源专有名词、2023-2024年度开源榜单汇编、开放科学和开放标准、开源合规最佳案例以附录形式进行补充收编，方便大家查阅。

我们希望《2024中国开源发展现状》能够为广大开源从业者、爱好者、用户，以及开源生态建设者、参与者，全面展现当前中国开源发展的全景图。限于时间及人力，其中定有不完善或疏漏之处，也请大家予以指正。

2024 中国开源发展的机遇与特色

政府推动的开源战略和政策环境，为中国开源发展带来了前所未有的机遇

中国政府高度重视开源发展，自开源被正式纳入国家“十四五”规划后，中国开源获得了前所未有的政策层面的大力支持。各地政府相继出台一系列政策和措施，切实推动当地开源发展，特别是国家大力推进科技自立自强和建设数字中国等战略，为中国开源产业的快速发展提供了良好的政策环境和市场机遇。

2023年，工信部提出中国开源立足新发展阶段，加快推进开源体系建设，释放开源发展潜能，推动高质量发展。打造开源应用示范标杆，支持国内外优质开源社区加快孵化和应用推广，助力提升关键软件源头创新和供给能力。来自政府的大力引导和支持，给中国开源发展带来前所未有的机遇。

中国开源最大的活力-开发者增长速度令人瞩目

- 中国开发者（含香港）在GitHub全球用户数量规模排名第三，2023年为1,100万，预计2026年GitHub中国开发者规模1,800 - 2,000万；活跃率居全球第一，显示了中国在开源领域的人才储备和发展速度令人瞩目。
- 开源新生力量偏向年轻化，增长空间和潜力巨大。中国开发者社区数据显示，中国开发者注册用户数4400万，2023年新增用户783万。随着开源的普及，很多年轻开源新生军逐渐加入和成长，开源人才增长空间和潜力巨大。
- 全国超过40%的开发者工作在一线城市（其中以北京、上海、广州、深圳为主）。北京、广东是开源开发者聚集较多的地域，占全国总数27%。上海、江苏地区占比数量处于第二梯队，占全国总数14%。
- 在开源项目的贡献上，开发者主要还是通过代码和文档的贡献，在参与过开源项目的开发者中，其中35%的开发者有代码贡献，31%有文档贡献。
- 全球开源大背景下，超过75%的开发者支持AI大模型开源，并认为这种开源模式可以促进全球开源社区参与和贡献。

开源项目稳定发展，部分中国企业表现抢眼

- 据GitHub统计，经过2018-2020年的高增长后，GitHub平台的年度事件增长数量逐渐下降。2023年GitHub日志数据为14亿，相比2022年增长了约10.32%，继续凸显出开源科技的发展在全球数字化转型中的活跃与关键地位。
- 中国大企业开源的项目体量及活跃度总量，与美国大企业相比，还是有不少的差距。放眼全球重点参与开源的企业，排名前五的分别是微软、华为、谷歌、亚马逊和Meta。
- 全球开源排行榜中，聚焦国内，华为于2019年将开源作为重要战略方向开始发力，仅用2年的时间就做到了国内第一，全球第二的位置；阿里巴巴在2021年前，一直是国内开源的领先者，并至今稳居世界第六的排名；蚂蚁集团于2023年正式进入世界前十；国内开源的第四大巨头百度，目前排名全球12。进入全球前30的中国企业还有字节跳动（18）、PingCAP（19）、飞致云（24）、Deepin（25）、腾讯（26）、以及乐鑫科技（27）。
- 聚焦中国开源项目，OpenRank排名位列前五的分别为openHarmony、openEuler、PaddlePaddle、MindSpore、Ant Design。企业与个人对于开源技术的贡献与参与度不断提升。

开源在重点领域的技术与应用，正逐渐强化和深化

- 开源已成为中国发展操作系统产业的重要途径。调研显示，中国操作系统发行版数量显著增长，尤其在物联网、嵌入式系统和智能网联汽车等领域。openHarmony正成为主流选择之一，特别是在物联网类别中，技术生态和影响力有显著提升。云和服务操作系统，openEuler、OpenAnolis和OpenCloudOS作为核心基础设施，持续领跑市场。
- 开源数据库在金融、电信、电商、医疗、物联网等行业中的应用案例不断增多，尤其是时序数据库如loTDB、TDengine、InfluxDB等，在物联网和工业互联网领域展现了强大潜力。
- 中国在大模型领域的开源活动相当活跃。阿里通义千问、智谱chatGLM、幻方量化deepseek-67B模型、零一万物的Yi-34B模型、百川的baichuan2-13B模型等陆续开源。中国企业在AI大模型的研发和应用上取得了显著成就，形成了竞争激烈的市场环境。
- 中国企业在RISC-V国际基金会中占有相当的比重，中国制造的RISC-V处理器占全球产量的一半。在RISC-V国际基金会高级会员中，中方成员占比达到48%；中国有超过400家企业关注、参与和使用RISC-V指令集进行开发。北京、上海、深圳、武汉、杭州等城市陆续推出一系列科研、产业化支持政策。

- 2023年中国汽车软件生态正在形成强大的开源生态：软件相关开源公司（如百度、普华基础软件、国科础石等）；代表性开源项目（如Apollo、DDS、ROS等）；开源组织（中汽协、汽车工程学会、OpenSDV汽车开源联盟）、汽车企业开源组织形态（如极氪汽车、长安、小米等OSPO）等蓬勃发展。

开源生态持续壮大，国际合作日益频繁

- 中国开源社区的建设和活动日益活跃，社区规模不断扩大，开源项目的数量爆发增长。2023年，由开放原子开源基金会牵头、阿里云、CSDN等企业共同建设的Atomgit代码协作平台上线。
- 中国开源社区和企业正积极融入全球开源生态，与国际开源社区的交流合作日益频繁。2023年北京国际开源社区正式启航；中国与欧洲各国、中东沙特阿拉伯的开源合作正迎来新的发展机遇。2024年3月，由智源研究院发起、中外顶级AI专家30多位共同签署《北京AI安全国际共识》。
- 中国在开源代码和项目的贡献上正以迅猛的速度增长，根据CNCF云原生计算基金会CTO Chris Aniszczyk介绍，中国已经成为第二大开源项目贡献国。截至目前，中国贡献了Harbor、Volcano、Dragonfly等32个项目，大约占CNCF所有项目的四分之一。
- 中国开发者已经成为Kernel.org社区最大贡献群体在过去的19年间，参与Kernel.org社区的中国开发者的规模提升了40倍，对社区的贡献绝对数量提升了50倍，对Kernel.org的贡献排名近五年来保持世界第一。
- ASF全球有800余位成员，其中，姜宁在2022年3月成功当选新任董事，2023年连任。伴随着越来越多来自于中国的项目进入到ASF，中国的Committer数量也在不断增加。自2014年开始，中国Committer数量逐年递增，直至今日，中国Committer数量占到新增Committer一半左右的比例。

开源教育纵横连合协同发展，人才培养规模化稳步推进

- 国内众多高等院校加大开源基础设施投入力度，进行开源课程建设，组织开源相关讲座、竞赛以及多种线上、线下活动，积极探索开源创新人才培养路径，提升软件人才与关键软件技术创新和供给能力。
- 知名开源项目及开源企业纷纷组建专门的技术培训学院，发布开源培训认证课程，围绕开源项目推

动专业化教育，如百度AI Studio、PingCAP TiDB培训及认证课程等。

- 开源教育正加速融入并贯穿教育全阶段：开源逐渐渗入各地高、中、小学教育，像Arduino、Scratch等开源软硬件在儿童创客教育产品中被广泛使用，众多开源组织和开源在线教育平台组织师资培训及学生创客活动，支持学校进行开源启蒙教育。同时，越来越多的职业教育机构和培训机构，大量引入和增设开源软件和工具、开源相关培训及网络课程。

- 政产学研用，多方协同，从开源文化、开源素质、开源技能“三位一体”协同培养，构建开源教育的生态发展。2023年中国计算机教育大会上，由教育部、CCF开源发展委员会、软件行业协会、开源基金会等多方一起，探讨大规模计算机实践教育资源共建共享方法及流程，支持高校加速推进计算机开源人才培养体系建设，从规模化人才培养方面赋能发展。

中国开源发展的风险和挑战

- 开源供应链安全问题。随着开源在各行各业的广泛应用，开源供应链的安全问题日益凸显。安全漏洞的持续增长、漏洞层级传播和修复滞后等问题，都可能给使用开源的企业带来巨大的安全风险。特别是在关键基础设施和敏感领域，一旦开源组件出现安全问题，可能会对整个系统的安全性造成严重影响。

- 开源的法律风险。开源许可证的多样性和复杂性给企业使用开源带来了法律风险。不同许可证之间的冲突、特殊附加条款的隐藏红线以及知识产权问题，都可能导致企业在使用开源时面临法律诉讼和经济损失。此外，由于开源的免担保性质，企业在使用过程中需要更加谨慎，避免因忽视风险而遭受损失。

- 技术创新风险和生态垄断。尽管中国开源社区和开源项目在数量上有所增长，但在质量和技术水平上与国际先进水平仍存在一定差距。许多关键技术和核心组件仍然依赖国外开源项目，这在一定程度上限制了中国开源产业的发展。同时，大型企业通过开源项目形成的生态垄断也可能阻碍技术创新和市场的公平竞争。这些风险需要通过建立更加健康和平衡的开源生态系统来解决。

- 人工智能热潮与五大风险。人工智能开源的确促进了技术进步和行业生态的繁荣，但也面临着安全漏洞与滥用风险、知识产权与商业利益冲突、模型偏见与伦理问题、合规与监管挑战、标准化与互操作性、可持续发展与社区维护等方面的问题。

- 国际形势和政治因素的影响。在当前的国际形势下，政治因素可能对开源合作产生影响。例如，出口管制和政治立场对立可能导致的开源项目分裂，这些都给开源的跨国合作带来了不确定性和挑战。

中国开源发展建议

中国开源发展正处于一个重要的机遇期，但同时也面临着不少风险和挑战。通过加强重点项目与领域的建设、推动生态共建、人才培养，提高安全意识、建立有效的组织和社区，深化国际交流合作，以及加大对重点开源项目的孵化，中国的开源事业将迎来更加可持续的发展。

继续加大对重点项目重点领域的支持

通过国家资金和政策支持，引导企业、学术界协同开展开源技术研究，对重点开源项目进行孵化，推动开源生态全面发展。同时，结合市场需求和行业特点，深入挖掘开源技术的商业价值，推动开源技术的广泛应用和产业化发展，为企业的可持续发展注入新的动力。

- 继续深入解决缺芯少魂的短板（解决操作系统和其生态芯片问题）；
- 继续加大开源深度信息技术发展（包括移动互联网、物联网、工业互联网、云原生、区块链、5G/6G通信、人工智能等）；
- 开发、维护、调整供应链（实现开源化、数字化），保障改善供应链／网络的安全运行。

加强开源生态的共建

为强化社会化组织和开源社区的建设，鼓励开发者和开发者社区普及开源文化、倡导开源精神，提升公众对开源的认识，激励和吸引更多开发者参与开源；建设好开源社区、开源代码托管平台、开源基金会等开源组织，健全社区治理机制，提高开源社区与开源项目的治理和运营能力；促进资源共享和技术交流，加强社区内部协作和知识共享；同时，鼓励企业支持和投资开源项目，为社区提供必要的资源和工具，以及通过政策引导和激励措施，保障开源社区的可持续发展和创新活力。

当前，开源商业化依然面临着极大的挑战，开源项目的贡献方、拥有方、开源社区力量、基金会或者OSI等行业非政府组织以及政府最终相互结合起来，形成了对于整个开源世界的运作和治理控制体系及至关重要的奖惩执行。

强化开源人才培养服务

建议进一步加强对开源人才的培养服务，包括加强国际顶级开源人才交流，推动中国开源人才高地建设；打造连接企业、行业 and 教育的“桥梁式”培养体系，设计以工作实践技能为基础的培养与认证体系，加强产教融合；同时推动开源成果在高校及业界的课程与认证体系设计、师资队伍建设和培养计划，培养开源创新人才，以支撑开源软件形成可持续发展生态。

提高对开源供应链安全的重视

为确保中国开源供应链的安全，建议加强安全意识教育，建立安全审查和漏洞管理机制，确保开源许可证合规性，制定项目评估体系，推动标准化和最佳实践，建立风险信息共享平台，鼓励社区参与，加强国际合作，制定应急响应计划，支持关键项目维护，探索开源保险机制，加强立法监管，提升供应链透明度，并培养专业安全人才，以此提高开源的安全性和可靠性。

加强国际合作与交流

鼓励企业、高校和科研院所在遵守国际开源规则的基础上，从国际开源底层平台切入开展协同研究，打造更多适用的应用场景，逐渐打破技术垄断。同时，跟踪研判源代码和建构在此基础上的软件/硬件适用的出口管制规则、法律规范，提高中国在开源领域的国际竞争力。

正如COPU主席陆首群所言，在国际形势愈发复杂的背景下，中国开源发展的核心是自主创新。中国开源社区要加强原创性的技术研发，推动自主知识产权发展，以实现关键技术领域的自主可控。中国开源人和开源组织更应该不忘初心、坚守本心，坚守并践行中立原则的开源发展之路，推动中国科技及全球开源的大发展。

第一章 开源通识

1.1 开源的定义和特征

1.1.1 开源的定义

开源 (Open Source) 一词于1998年2月3日由Chris Peterson提出, “Open Source” 的概念出自当时著名的黑客 (Hacker) 社区Debian的社长Bruce Perens起草的“Debian Free Software Guidelines” (DFSG)。1998年2月下旬, Eric Raymond和Bruce Perens共同创立"Open Source Initiative" (OSI)。OSI承担的首要任务之一是起草开源定义 (OSD), 并使用它来开始创建OSI批准的许可证列表。

对确立“Open Source”定义有决定意义的是在1998年4月7日在美国加州Palo Alto由18位“自由软件运动领袖”召开的Freeware高层会议, 通过了传播开源的必要性。会议由Tim O'Reilly主持, Brian Behlendorf (Apache创始人)、Linus Torvalds (Linux创始人)、Guido Van Rossum (Python创始人)、Eric Raymond (著名记者、OSI首届主席) 等参加。

随着时代的变迁, 开源的定义内涵也在不断演进。目前的共识是: 开源是促进信息技术创新的重要途径, 是将源代码、设计文档或其他创作内容开放共享的一种技术开发和发行模式。技术领域的开源已包含源代码+软硬件设计文档源代码、源数据等技术和资源的开放共享, 源代码仍然是开源的主要内容。在开源模式下, 通过许可证的方式, 使用者在遵守许可限制的条件下, 可自由获取源代码等, 并可使用、复制、修改和再发布。本质上, 开源模式是一种依托互联网平台, 大规模群体智慧通过共同参与和协作, 不断累积智慧, 实现持续创新的方法。该方法中, 项目的核心开发人员与大规模的外围群体紧密合作, 他们通过互联网来共享资源、开展协同开发、管理代码等, 由此使得项目开发的效率、应对需求变化的能力大幅提升。

1.1.2 开源的特征

COPU于2016年归纳提出的 (已被国际认可) 开源特征包括: 开放、创新、自由、共享、协同、绿色、民主 (化)。开源的本质在于开放、共享、协同。

- 开放指开放标准、开放环境、开放源码；
- 创新是主轴；
- 自由指自由发布、自由传播、自由复制、自由修改、自由使用，其中，自由传播是开源运动的要义；
- 共享指共享资源；
- 协同指协同开发、协同作业、协作生产；
- 绿色指支持绿色可再生能源、绿色环境和零边际成本效应；
- 民主化指在新兴协同共享中，创新和创造力的民主化正在孵化一种新的激励机制，这种机制很少基于经济回报，而更多地基于推动人类的经济生活方式，缩小收入差距，实现全球民主化。为避免原创技术的流失，不能单纯限制孵化阶段开源代码的自由传播，可通过开源的商业模式、安全模块、运维举措、生态系统所构筑的屏蔽层来解决。

作为一种创新协作模式，开源已经不仅仅是开放源代码的软件技术开发，还包括更为广泛的开放技术领域及协同创新的理念与机制，总体包括开放科学、开源软件、开源硬件、开源技术、开源文化、开源经济等。“互联网+基于知识社会的创新2.0”是开源创新的基础理论（2005年由COPU提出，后来获得全球开源界领袖们的确认），与工业4.0、工业互联网机制相通，与元宇宙（Metaverse）概念相似。利用开源渠道，将高阶社会（知识社会，虚拟化的实验空间）中的技术、管理、资源等要素，作用于低阶社会（现实的工业社会，物理空间）中的业态（生产的、技术的、经济的、社会的），促使其产生0→1的爆发性重构。

1.1.3 开源的范畴

时至今日，开源已成为开源软件、开源硬件、开源生态、开源技术、开源社区、开源经济、开源商业模式、开源理念、开源文化、开源教育、开源许可证、开源基金会、开源孵化器、开源数字化治理体系、开源标准等的总称。

1.2 世界开源发展简史

世界开源发展史，有三个里程碑时间节点：1970年、1985年、1991年，分别对应UNIX、GNU、Linux的发展。

1.2.1 UNIX简史

1969年之前，Bell实验室和MIT、GE合作启动开发一个多任务多用户的操作系统Multics，为UNICS操作系统诞生创造条件。

1969年，UNIX原型——UNICS（UNiplexed Information and Computing System）诞生。

1973年，Dennis Ritchie发明C语言，用C语言将UNICS重写并更名为UNIX，UNIX正式诞生。

1969-1977年，UNIX相继推出V1-V6版本，向社会开放源代码。

1977年开始进入“后UNIX”时代，UNIX演化为UNIX（闭源，如AT&T-UNIX等）和BSD（开源，如BSD-UNIX等）。主流的开源BSD操作系统有386BSD、FreeBSD、NetBSD、OpenBSD。

因此，世界开源的历史始自1970年开启的“前UNIX”，至今53年，开源在事实上已经历了半个世纪的发展历程。

1.2.2 GNU简史

GNU系统是一套向上兼容UNIX的完全自由的操作系统。GNU代表“GNU's Not UNIX”。

1983年9月27日，Richard Stallman提出了GNU工程的初始声明，意图找回在早期计算机社区中广泛存在的合作精神，通过除去专属软件所有者设置的障碍而使得合作再次成为可能。

1984年1月，Richard Stallman从MIT离职并开始编写GNU软件，同年9月开始GNU Emacs的工作。

1985年，Richard Stallman发表《GNU宣言》，吸收“前UNIX”和BSD的开源成果，开发推出Emacs等编译器等自由软件（Free Software）。同年10月，自由软件基金会（FSF）正式成立。“Free Software”中的“Free”强调的是自由，而不是免费，是指可以付费或不付费得到GNU软件。一旦得到了软件，便拥有了使用它的四项特定自由。可以自由按照自己的意愿运行该软件；可以自由复制软件并将其送给朋友和同事；可以自由通过对源代码的完全控制而改进程序；可以自由发布改进的版本从而帮助社区建设。

1.2.3 Linux简史

Linux的诞生、发展和成长过程始终依赖着五个重要支柱：UNIX操作系统、MINIX操作系统、GNU计划、POSIX标准和Internet网络。

1977年，UNIX的重要分支——BSD在UC Berkeley诞生。开发者是SUN公司创办者Bill Joy。

1984年，Andrew S. Tanenbaum为能继续进行学术教育研究，重新开发了一个基于“前UNIX”和BSD4.3的开放源代码操作系统——Minix，这是Linux最开始的前身。

1985年，GNU开源计划和FSF基金会成立。在这个时间前后，涌现了很多重要的软件和协议，如GPL协议、GCC编译器、Emacs编辑器、Glibc、Bash shell等，为Linux后来的发展创造了肥沃的土壤。

1991年，芬兰本科生Linus Torvalds在Minix和GNU成果的基础上，开发并以GNU GPL许可方式发布了开放源代码操作系统Linux v0.01。

1992年5月21日，Peter MacDonald发布第一个独立的Linux安装包SLS。

1993年6月17日，Slackware Linux由Patrick Volkerding发布，这是第一个取得广泛成功的Linux发行版。8月16日，Ian Murdock（Debian中的“ian”）发布了第一个Debian Linux发行版。Debian是最具影响力的Linux发行版之一，是MEPIS、Mint、Ubuntu等的鼻祖。8月19日，Matt Welsh写的《Linux Installation and Getting Started》第1版出版，这是第一本关于Linux的书籍。

1994年3月14日，Linux内核V1.0发布，它支持基于i386单处理器的计算机系统。8月15日，William R. Della Croce, Jr.申请了“Linux”商标，9月进行了注册。11月3日，Red Hat共同创始人Marc Ewing宣布可以以49.95美元的零售价格获得Red Hat Software Linux的CD-ROM和30天的安装支持。

1996年5月9日，吉祥物Tux诞生。6月9日，Linux内核V2.0发布，这是第一个在单系统中支持多处理器的稳定内核版本。Linux从此成为很多公司的重要选择。10月14日，Mattias Ettrich发起KDE项目，此前UNIX和Linux都没有统一的桌面系统，编写桌面软件非常复杂。

1997年1月9日，第一个“Linux病毒”Bliss被发现。

1998年5月1日，基于Linux的Google搜索引擎面世。12月4日，IDC报告称1998年Linux出货量至少上升200%，以其他任何操作系统无法企及的速度增长着。

1999年3月3日，GNOME桌面系统进入Linux世界。

2000年2月4日，IDC报告表明Linux排在“最受欢迎的服务器操作系统的第2位”。10月30日，第一个Linux live发行版Knoppix发布。

2001年1月，Linux 2.4发布，进一步提升SMP系统的扩展性，同时集成了很多用于支持桌面系统的特

性: USB、PC卡 (PCMCIA) 的支持, 内置的即插即用等功能。

2003年3月6日, SCO Group公司宣布正在发起对IBM高达10亿美元的诉讼, 声称IBM把SCO的商业机密整合到了Linux中。12月, Linux 2.6版内核发布。

2007年8月8日, Linux基金会由开源发展实验室 (OSDL) 和自由标准组织 (FSG) 联合成立, 得到了主要的Linux和开源公司, 包括富士通、HP、IBM、Intel、NEC、Oracle、Qualcomm、三星以及开发者支持。11月5日, Google发布Android, 被称为“第一个真正开放的综合移动设备平台”。

2011年5月11日, Google I/O大会发布基于Linux内核的云操作系统Chrome OS的Chromebook。6月21日, Linus Torvalds发布Linux3.0版本。

Linux内核大版本的更新周期一般是2-3个月, 2023年5月24日, 稳定版本Linux 6.3.4, 发布, 其Maintainer为Greg Kroah-Hartman & Sasha Levin。2024年7月14日, Linux Kernel 6.10版本发布。

1.2.4 开源基金会崛起

开源基金会是开源生态中的重要部分。对于非营利性且处于中立位置的开源基金会来说, 他们拥有开源项目的知识产权, 而且没有任何商业利益, 这为贡献者、开发者、用户提供了良好的协作平台;基金会也会通过技术服务和项目孵化来帮助开发者和开源企业, 如软件仓库、问题跟踪、技术指导、法律支持、项目投资、公共关系维护等等;基金会还会提供项目日常的运营和治理, 以满足在项目生命周期不同阶段对于管理项目的需求。自20世纪80年代自由软件基金会(Free Software Foundation)开始, 全球各类开源基金会快速发展, 在构建全球开源生态中发挥着巨大的促进作用, 据不完全统计, 各类开源基金会已经超过100家, 其中具有代表性的包括(以下数据截至2024年7月):

自由软件基金会 (1985年成立)。通过促进自由软件和文档 (特别是GNU操作系统) 的开发和使用。以及通过反对数字限制管理 (DRM) 和软件专利对计算机用户自由的威胁。致力于确保计算机用户的自由。

国际互联网工程任务组。它是互联网的首要标准开发组织 (SDO), 主要任务是负责互联网相关技术规范的研发和制定。目前, 它已成为全球互联网界最具权威的大型技术研究组织。

国际互联网协会 (1992年成立)。致力于赋予人们力量, 使互联网成为一种积极的力量: 开放、全球连接、安全和值得信赖。

万维网联盟（1994年成立）。通过开发协议和指南，确保了Web的长期增长，引领Web发挥其最大潜能。

公共利益软件组织（1997年成立）。一个位于美国的非营利组织，旨在帮助组织开发和分发自由硬件和软件。

开放源码促进会（1998年成立）。致力于推广开源软件发展。

Apache软件基金会（1999年成立）。为开源软件项目提供支持。

GNOME基金会（2000年成立）。是一个非营利性组织，致力于通过建立一个多样化且可持续的自由软件个人计算生态系统。

Open Source开发实验室（2000年成立）。是一个非盈利组织，其目的是促进Linux及其相关技术的增长和采用。

FreeBSD基金会（2001年成立）。是一个基于美国501(c)条款注册的一家非营利组织，致力于支持FreeBSD项目的发展和社区。

Python基金会（2001年成立）。致力于推进Python编程语言相关的开源技术的组织。

Blender基金会（2002年成立）。是一个独立的公益组织。旨在提供一个完整、免费和开源的3D创作管道。

Mozilla基金会（2003年成立）。目标是保证Mozilla可以在没有网景以后能继续生存下去。

Eclipse基金会（2004年成立）。一个国际非营利性协会，提供了一个成熟、可扩展且对业务友好的环境，用于开源软件的协作和创新。

Open invention Network（2005年成立）。旨在消除核心开源软件技术遇到的专利风险，推动实现更高的技术创新和行业发展。

软件自由保护组织（2006年成立）。旨在为自由开源软件项目提供支持和基础设施的非营利组织。

Linux基金会（2007年成立）。一个非盈利性的联盟，由OSDL和FSG合并而成，其目的在于协调和推动Linux系统的发展，以及宣传、保护和规范Linux。

OpenStack基金会（2010年成立）。一个专注于开源云计算项目的非营利组织，旨在开发、支持、保

护和推广开源基础设施。2020年10月26日，OpenStack基金会已经正式更名为开源基础设施基金会（Open Infrastructure Foundation, OIF）

云原生计算基金会（2015年成立）。致力于推广云原生计算采用的开源软件基金会。

RISC-V基金会（2015年成立）。负责定义RISC-V作为开放标准指令集架构的国际组织。

学院软件基金会（2018年成立）。为电影和更广泛的媒体行业的开源软件开发人员提供一个中立的论坛。

OpenJS基金会（2019年成立）。致力于推动JavaScript和开源技术的发展。

CD FOUNDATION持续交付基金会（2019年成立）。持续交付的行业规范。

TARS基金会（2020年成立）。它是Linux基金会旗下的一个非盈利性开源微服务基金会。它致力于构建一个健康、开放的微服务开源生态，让企业能在创新领域和扩展应用程序时充分利用微服务架构的优势。

SODA基金会（2020年成立）。它是Linux Foundation旗下的一个开源项目，旨在培养一个开源数据管理和存储软件的生态系统，以实现数据自治。

开放原子开源基金会（2020年成立）。它是致力于推动全球开源事业发展的非营利机构。开放原子开源基金会专注于开源项目的推广传播、法务协助、资金支持、技术支撑及开放治理等公益性事业，促进、保护、推广开源软件的发展与应用;致力于推进开源项目、开源生态的繁荣和可持续发展，提升我国对全球开源事业的贡献。

OpenSSF基金会（2020年成立）。旨在推动开源软件的安全性发展。

Rust基金会（2021年成立）。一个独立的非营利组织，致力于管理 Rust 编程语言，培育 Rust 生态系统，并支持管理和开发项目的维护者。

天工开物开源基金会（2023年成立）。于2023年3月获得重庆市政府批复成立，由清华大学，华中科技大学，重庆大学，中冶赛迪，中国信科，中科院等知名高校和企业联合发起，致力于推动中国开源产业公益事业的非营利性独立法人机构。

1.3 中国开源发展简史

1991年，中国与AT&T Bell Laboratories USL/USG合作，引进UNIX SVR4.2版本源代码(全球唯一，但属于“后UNIX”闭源时期)，并发布了中文版本合作组建了中国UNIX公司;中方同时也引进“前UNIX”开放的源代码。

1999年，中科红旗、中软网络、冲浪平台在引进Red Hat公司Linux发行版的基础上，分别推出最早的Linux中文版本。

2008年，以Linux基金会为代表的一众国际开源基金会陆续来华举办活动设立分支机构，是国际开源生态全面认可中国的开源贡献，接纳中国开发者和中国企事业高校科研机构的一个重要转折点。

2019年，中国开源贡献在国际开源社区占比显著提升，一些领域已经处于相对领先地位。在这一时期，中国开源生态得到了全面发展，一些中国企业和开发者逐步进入国际开源基金会的核心管理团队，源自中国的开源许可证、开源代码托管平台崭露头角；在自动驾驶、机器学习、人工智能等领域的中国项目在国际上引起很大反响。中国开源生态的发展已经呈现出“由点及面”的重要转折。

因此，中国开源的发展历史具有四个里程碑节点，1991年、1999年、2008年、2019年，可以代表着中国开源进入启蒙、萌芽、发展、加速四个时代的里程碑。

1.3.1 启蒙时代

1991年，陆首群、张克治和杨天行牵头代表中方代表团与美国AT&T USG合作，美方将最新开发的UNIX版本——UNIX SVR4.2源代码向中方开放。中方为此组织了UNIX新版本编辑委员会，由杨芙清、胡道元、仲萃豪、刘锦德、尤晋元、贾耀良、孙玉芳等国内资深软件专家组成，并邀集全国软件专家、程序员200多人，翻译、编辑、出版了UNIX SVR4.2中文版本共19册，首发式在人民大会堂举行，向全国各大图书馆、高等院校赠书，业界影响甚大。这一历史事件被Linux基金会收录，后被作为中国开源诞生的标志性事件。1991年成为中国开源元年。

1992年12月，中方与USG合资在华成立了中国UNIX公司。

1994年，已经在芬兰工作5年的宫敏博士第一次将Linux和大量自由/开源软件源码通过磁带带回中国。

1997年，宫敏第二次带回80GB的自由软件，并在国家信息中心的帮助下在中国经济信息网上建成了“中国自由软件库”，为国内技术人员了解、学习和使用Linux及相关自由/开源软件提供了便利。

1998年，在开源概念提出的这一年，在湖南长沙读博士的章文嵩在Linux2.0内核上利用课余时间开发

了Linux虚拟服务系统（LVS），并在第一时间开源，在全世界引起很大反响。同一年，在北京清华大学读研的魏永明，决定开发针对实时嵌入式系统的图形界面系统并将其开源，MiniGui应运而生。同一年，在安徽合肥读大二的自动化系学生吴峰光开始接触Linux，他发现Linux的文件预读算法磁盘I/O性能较差，便尝试对Linux内核进行改进。

1998年初，北京冲浪平台软件技术有限公司成立，专注于研发Linux中文发行版，同年9月Xteam Linux收获了第一个市场用户——北京市政府政策研究室。1999年3月正式对外发布Xteam Linux中文版1.0。2001年冲浪平台在香港创业板上市。

1.3.2 萌芽时代

在历时8年的启蒙阶段中，自由/开源软件的理念和文化开始在中国大地生根发芽，随着UNIX、Linux在中国用户的激增，以及互联网时代的到来，中国自发地涌现出一批开源社区和开源企业。在中国开源发展的萌芽阶段，Linux技术的快速普及起到了核心的推动作用。

1999年9月，蓝点软件技术（深圳）有限公司成立，次年3月在美国纳斯达克上市。一个中国初创公司从成立到在美上市仅用不足200天，且上市当天股价暴涨400余倍，这个神话旷古绝今。

1999年底，时任中国科学院软件所副所长的孙玉芳老师牵头筹建北京中科红旗软件有限公司和北京红旗中文贰仟有限公司，分别基于Linux和OpenOffice进行中文版红旗Linux和RedOffice的研制。孙玉芳老师英年早逝（2005），一生发表操作系统相关论文180余篇，编著和译著操作系统原理相关著作140余册，他指导过的诸多硕士/博士生现已成为开源产业的中坚力量。1999年，中国软件开发者社区CSDN在蒋涛的努力下成立了，并于2000年创办《程序员》杂志。

2000年2月，在科技部国家高技术研究发展计划（863计划）的支持下，由一批国内高校、科研院所、IT企业联合发起成立了共创软件联盟。共创软件联盟在开源许可证规则下对863计划的软件成果进行开源培育和孵化，并提出以开放源代码的协同创新模式作为中国基础软件发展的主要路径，对我国开源产业发展起到了重要的推动作用。

2002年，黄建忠以CJacker的ID在中国Linux公社发布了基于RedHat 8再发布的Magic Linux版本——这是国内第一个Linux社区发行版。在这个阶段，部分中国开发者开始尝试向国际上游社区贡献代码。但同时，国内Linux企业对开源技术采用拿来主义，忽视开源社区、商业模式和开源知识产权等方面的投入和积累，导致在国际社区中出现了许多针对中国Linux发行商只是开源使用者的评价，直

到2008年，这种印象才得以扭转。

2003年，原信息产业部通过电子发展基金支持“Linux公共开发平台”的建设，并成立国家软件与集成电路公共服务平台，组织国内操作系统厂商进行Linux参考平台的研发和Linux操作系统的应用推广。

2004年，COPU成立，陆首群教授被推举为联盟主席。在陆首群主席的大力推进下，COPU为推动中国开源软件的发展而不懈努力，也为促进中、日、韩乃至全球关于开源运动的沟通、交流、共享、协同与合作而努力，为促进中国、东北亚和全球开源运动作出贡献而努力。联盟的国际智囊团聘请了几十位全球顶级开源专家以多种形式参与和推进中国开源建设，如此规模的智囊团在全球首屈一指，是开源界公认的顶级智库，是COPU创建的开源高地、科技高地、创新高地、人才高地。可以说，联盟的成立和陆主席的领导加速了中国开源软件发展的速度，为碎片化的国内开源产业建立了统一的根据地，为国际开源生态和中国开源生态建立了纽带。

2008年，以Kernel、Apache、GNOME、OpenOffice为代表的国际知名开源社区纷纷选择在北京召开全球技术峰会，作为他们进入亚洲的第一步（之前从未在亚洲国家召开顶级技术峰会）。这是对中国开发者在社区提交贡献的肯定。因此，我们认为2008年成了中国开源发展重要的分水岭。

1.3.3 发展时代

在2009年前，中国常被视为国际开源资源单纯的使用者，2009年之后，中国开发者提交的代码被Linux基金会接受，排名稳居世界前四（Kernel 2.6.27之后），中国开始成为国际开源资源的重要贡献者。

在2009年之后，伴随着云计算、物联网、大数据、移动互联网的快速发展，中国一批有实力有抱负的知名科技IT企业、互联网公司开始选择拥抱开源。Linux基金会执行董事Jim Zemlin曾表示：“这些中国科技公司无论从开源技术上，还是产业、经济发展上，已经走进世界领导者行列。”

在2009-2018年这十年中，中国开源呈现了百花齐放的繁荣局面，从开源开发者数量的激增，到开源社区的多样化（不再局限Linux等操作系统领域），再到知名科技企业自内而外的拥抱开源（内部实行开源治理，逐渐对外捐献开源项目，并向上游技术社区提交贡献）。2012年成立的中国开源云联盟（COSCL）和2016年成立的云计算开源产业联盟（OSCAR）就是典型的开源与云计算技术融合的产业联盟。这一阶段明显的特点在于不论是初创企业还是千百亿体量的科技巨头纷纷用开源技术来构

建公司自身的技术堆栈，在技术上全面拥抱开源，并越来越重视对上游开源社区和开源基金会的投入与贡献。

1.3.4 加速时代

自2019年起，中国开源进入了一个新的时代——加速阶段。

根据Linux基金会有关Kernel.org社区贡献的统计，来自中国开发者的贡献已居全世界之首（Kernel 4.14至今）。中国开源发展进入深水区，开源生态的完整性、开源社区的成熟性、开源商业模式的多样性和复合性都在发生巨大变化。

由木兰开源社区牵头、多家单位联合参与起草的“中国第一个开源协议——木兰许可证”已经正式发布，前不久还被OSGJ翻译为日文版使用。2020年6月，开放原子开源基金会在北京正式成立，是我国唯一的国家级开源基金会。

在现阶段，部分领先的中国企业不满足于拥抱开源，他们正在很多创新领域跃跃欲试，希望通过开源模式来对创新技术进行引领，如商汤的OpenMMLab计算机视觉开放算法、矩阵元的Rosetta隐私AI开源框架、阿里的RocketMQ海量信息中间件、京东的JD Chain开源区块链、百度的PaddlePaddle深度学习平台和Apollo自动驾驶平台、海尔卡奥斯COSMOPlat工业互联网平台的工业APP开发框架开源项目等。同时，在现阶段出现了一些原生的开源明星项目，如Apache SkyWalking、Apache Echarts、Apache IoTDB、TiDB，其中一些优秀的开源项目如TiDB备受资本市场的推崇，创造了全球数据库融资历史新的里程碑。

2023年，大模型引起了AI开源热潮。6月9日，北京智源人工智能研究院院长黄铁军在2023北京智源大会上作报告，发布悟道3.0大模型系列，并宣布进入全面开源的新阶段。6月15日，百川智能公司，推出了70亿参数量的中英文预训练大模型——baichuan-7B；7月，百川智能继续推出了baichuan-13B，不仅对学术研究完全开放，所有开发者均可通过邮件向百川智能申请授权，在获得官方商用许可后即可免费商用。7月14日，智谱AI官网发布：经智谱AI及清华KEG实验室决定，自即日起ChatGLM-6B和ChatGLM2-6B权重对学术研究完全开放，并且在完成企业登记获得授权后，允许免费商业使用。8月3日，阿里云开源通义千问70亿参数模型，包括通用模型Qwen-7B和对话模型Qwen-7B-Chat，两款模型均已上线魔搭社区，开源、免费、可商用。9月4日，既阿里达摩院推出的魔搭社区之后，始智AI团队发布始智AI-wisemodel社区正式上线，纷纷将目标指向中国版“HuggingFace”。10月30日，昆仑万维宣布开源百亿级大语言模型“天工”Skywork-13B系列，并罕

见地配套开源了600GB、150B Tokens的超大高质量开源中文数据集。11月22日，中国人民大学高瓴人工智能学院的大语言模型系列——YuLan开源。2024年1月10日，中国电信宣布，将自研星辰AI大型模型全面开源，以更好地助力客户业务的智能化升级。

2024年3月12日，北京大学深圳研究生院-免展智能AIGC联合实验室正式发布了“Open-Sora”计划旨在复现一个“开源版Sora”。该项目由深圳免展智能创始人、CEO董少灵与北京大学信息工程学院助理教授、博士生导师袁粒领衔发起，期望通过产业公司和科研机构的强强联手，以开源的方式对外传播中国团队的技术自信，号召全球社区一起做人类普惠的技术内容。

1.4 开源软件与知识产权

至此，我们对开源的定义和发展史有了基本的认识，在继续深入了解之前，有必要关注开源软件与知识产权的知识。

1.4.1 软件知识产权的主要内涵

软件知识产权是计算机软件人员对自己的研发成果依法享有的权利。目前国际上对软件知识产权的保护法律还不是很健全，大多数国家都是通过著作权法来保护软件知识产权的，与硬件相关密切的软件设计原理还可以申请专利保护。

- 软件知识产权主要包括著作权、专利权与商标权。
- 著作权指软件的表达（如程序代码、文档等）方面的权利。
- 专利权包括软件的技术设计，如程序设计方案、处理问题的方法、各项有关技术信息等方面的权利。
- 商标权则是指软件的名称标识方面的权利。

软件需要有硬件平台的支撑才可以运行，而在硬件平台上运行的软件基本上是目标码（二进制格式），从而造成软件源码和二进制目标码的分离和割裂，而软件作为通用商品进行销售或分发时主要提供的只是二进制文件，这也成为主流的软件销售和分发方式，并主要通过随软件一起发布的最终用户许可协议（End User License Agreement, EULA）对软件的使用、修改和分享及其他相关事宜作出规定。

针对上述对软件源代码的处置方式，出现了一种开放、协同、共享的软件发展模式——通过开源软

件许可证对软件源代码的复制、修改、再分发等权益，对商标、专利、著作权等内容进行进一步规范，这也是对软件知识产权发展的一个重要补充。

1.4.2 开源软件许可证

开源许可证就是开源软件使用的许可证。对开源软件许可证的法律界定先前存在一定的争议。在美国，一些法院认为软件许可证是合同（Contract），一些法院则认为是许可（License）。两者的区别在于，许可在传统上是由地产或物主作出的，目的在于允许他人使用自己的地块或物品。因此，它是单方向的，不构成完整的合同，而是作为合同的一个要素，用来和他人交换的条件。

合同和许可之分在美国法上具有重要意义。如果是合同，那么需要适用各州不同的合同法；如果是许可，那么需要适用统一的联邦版权法。此外，合同的违约救济和版权的侵权救济也有诸多不同，比如禁令（行为保全）的适用、判赔额的确定、律师费的分担等。

不过，包括中国在内的大陆法系国家，则普遍认为开源软件许可证构成合同；只不过这种许可合同并非协商得到，而是事先规定好的标准化格式合同。具体来讲，开源许可证是涉及版权、专利、商标等一系列权利义务的格式合同，且自动生效。

法律并没有限定许可证不能包含什么条款，这导致许可证的类型极其繁多、内容也非常自由。据不完全统计，广义上的开源许可证目前有超过200种，即便是OSI批准的许可证目前也多达96种，其中包括由中国主导编制的Mulan Permissive Software License v2(MulanPSL - 2.0)，全部许可证可参见<https://opensource.org/licenses/alphabetical>和<https://opensource.org/licenses/category>。

1.4.3 国外主流开源许可证解读

（1）GNU通用公共许可证（GPL）

GPL通用公共许可证是一个许可证家族的泛称，是一个被广泛使用的自由软件许可协议条款，为用户提供提供了足够的复制、分发、修改的权利：

- 可自由复制；
- 可自由分发，提供他人下载；
- 可用来盈利，分发软件过程中收费（必须在收费前向客户提供该软件的GNU GPL许可协议）；
- 可自由地修改（使用了这段代码的项目也必须使用GPL协议）。

GPL当前主要有GPL v2和GPL v3两个版本。这组许可证的共同特征在于其“传染性”：任何基于GPL代码编写的软件都必须成为开源软件。换言之，使用了任何GPL代码的软件，无论GPL代码占比多少，都必须将完整的源代码公开，并允许他人修改、发布。

GPLv2许可证存在一些漏洞，例如它不能阻止一个软硬件结合的系统中，通过对硬件部分施加限制，间接阻止用户在该硬件上运行软件的修改版本；没有包括关于专利的约定，导致实践中出现Microsoft-Novell专利协议这类试图将专利申请用作对付自由软件社群的武器的现象。

为解决这些问题，GPL v3于2007年发布。除填补上述漏洞，GPL v3兼容性更好。自由软件基金会明确表示GPL v3与Apache 2.0许可证兼容。目前，使用GPL许可证的重要项目包括Linux内核和MySQL等，但新兴项目一般会选用更宽松的许可证。

(2) LGPL (Lesser General Public License)

LGPL叫做GNU宽通用公共协议，对产品所保留的权利比GPL少，适用于非开源产品的开源类库或框架，并可以不继承LGPL协议，使用其他的开源协议。

LGPL的特点在于，链接到该软件库的软件可以不适用LGPL或GPL，换言之，可以不公开源代码。LGPL的这一特性消除了在GPL下软件商用的最大障碍。尽管如此，基于该库修改而得到的软件仍然需要遵循GPL许可证。

(3) Mozilla Public License (MPL)

MPL允许在其授权下的源代码与其他授权的文件进行混合，包括私有许可证。在MPL授权下的代码文件必须保持MPL授权，并且保持开源，即允许在派生项目中存在私有模块，同时保证核心文件的开源。使用MPL授权的软件并不受专利的限制，其可以自由使用、修改，并可自由的重新发布。带有专利代码的版本仍然可以使用、转让，甚至出售。

(4) BSD许可证系列

BSD许可证是由UC Berkeley首倡和维护的，版本繁多，目前常用版本包括原始的BSD-4、BSD-3以及简化的BSD-2。BSD属于宽松许可证，与MIT许可证接近但更加宽松，甚至跟公有领域更为接近。在最简化的BSD-2许可证下，保留著作权声明、许可证内容以及免责声明即可；只要满足许可证设定的条件，就可以自由地修改并发布代码。

BSD-3许可证在BSD-2许可证的基础上增加了禁止背书条款（未经事先书面许可不得使用原作者之名来推广衍生作品）；BSD许可证进一步增加了广告条款（衍生作品的广告材料必须说明该软件包含由UC Berkeley及其贡献者开发的软件）。

（5）MIT许可证

MIT许可证之名源自首倡者MIT麻省理工学院，又称“X许可协议”（X License）或“X11许可协议”（X11 License）。

据统计，2015年GitHub上高达45%的项目使用MIT许可证。近几年MIT许可证的份额有所下滑，在2020年市场份额第一的位置被Apache 2.0取代，但仍是最受开发者欢迎的许可证之一。

MIT的特点在于条款非常简单，是开源协议中最宽松的一个：“被许可人有权利使用、复制、修改、合并、出版发行、散布、再许可和/或贩售软件及软件的副本，及授予被供应人同等权利”，加上要求被许可人保持同样的声明。

（6）Apache许可证

Apache许可证是一种“宽松”（Permissive）的许可证，目前常用版本是2.0，2020年成为使用最多的许可证。“宽松”是指不保证被使用软件的派生版会继续保持自由软件的形式，即：“怎么用都行，用在哪都行”。具体而言，Apache 2.0许可要求保留版权和许可声明，但允许许可作品、修改和更大的作品在不同的条款和没有源代码的情况下分发，只是未修改的部分仍然需要保留Apache许可证。除了为用户提供版权许可之外，还有专利许可，对于那些涉及专利内容的开发者而言，该协议最适合。由于上述利好条件，Apache 2.0成为了相当多流行的开源项目的许可证，最著名的例子之一就是Kubernetes。

不同开源许可证就许可、条件和限制的比较参见<https://choosealicense.com/appendix/>。主要常见开源许可证的对比汇总见下表。

可以看出，除了允许商业性使用、分发和修改外等常见特征外，开源许可证还有两个共同点。一是开发者不承担保证责任（瑕疵担保责任）。开源代码通常都是免费提供的，因此开发者不应为他人使用该软件造成的损失而承担责任。二是要求保留著作权标记。开源软件并不意味着放弃著作权。相反，开源许可证的强制效力来源，恰恰来自于作者对开源软件的著作权。

因此，许可证一般都要求以适当的形式保留著作权标记（包括许可证正文以及作者署名）。

1.5 开源治理

许可证名称 \ 许可 条件 限制	商业用途	分发	修改	专利使用	私有使用	披露来源	许可和版权声明	网络使用是分布式	相同许可证	状态更改记录	责任	商标使用	保证
Apache许可证 V2.0	是	是	是	是	是		是			是	否	否	否
BSD-3 Clause Clear License	是	是	是	否	是		是				是		否
BSD-2 Clause Simplified License	是	是	是		是		是				否		否
GPL V2.0	是	是	是		是	是	是		是	是	否		否
GPL V3.0	是	是	是	是	是	是	是		是	是	否		否
LGPL V2.1	是	是	是		是	是	是		是?	是	否		否
LGPL V3.0	是	是	是	是	是	是	是		是?	是	否		否
AGPL V3.0	是	是	是	是	是	是	是	是	是	是	否		否
MIT许可证	是	是	是		是						否		否
Mozilla Public License V2.0	是	是	是	是	是	是	是		是?		否	否	否
Eclipse Public License V2.0	是	是	是	是	是	是	是		是		否		否
木兰宽松软件许可证 V2.0	是	是	是	是	是		是				否	否	否
开放数据共享开放数据库许可证	是	是	是	否	是	是	是		是		否	否	否
Microsoft Public License	是	是	是	是	是		是					否	否
SIL开放字体许可证 V1.1	是	是	是		是		是		是		否		否
知识共享署名4.0国际 (CC-4)	是	是	是	否	是		是			是	否	否	否

常见开源许可证的对比汇总

1.5.1 开源治理的概念

开源治理是指开源项目管理和运营的全过程。它涉及到开源战略、治理结构、决策机制、贡献指南、合规审查和开发者关系运营。有效的开源治理是在开放共享、协同创新的开源文化基础上，确保项目发展的持续性、决策流程的透明性和社区协作的有效性的，一组开源行为准则和相应的监督管理机制。以下是开源治理中通常包括的一些关键要素：

开源战略：对于开源基金会，开源战略基本等同于其成立宗旨；对于科技企业，开源战略则是其企业发展战略的重要组成部分。通常情况下，开源基金会的开源战略侧重于对开源项目的运营和孵化，而科技企业的开源战略着眼于开源合规和投资回报。

治理结构：较为常见的三种开源治理结构特征包括，仁慈独裁治理、精英治理和民主治理。在当代国

际主流开源社区中，特别是一些大中型开源社区，往往采用分布式的治理结构。如，对待非核心、非紧急性的事务更多采用民主治理模式；而对待核心的或紧急性的事务更多采用仁慈独裁和精英治理的治理结构。

决策机制：成熟的开源基金会通常具备上述三种典型治理结构中的一种或两种特征。而科技企业则具有中心化的组织特征，与企业管理体系保持一致，即使设有OSPO或技术管理委员会这样的机构，也同样趋近于仁慈的独裁治理模式，由CEO或CTO等企业最高决策机构来“集中式”决定。不同的开源社区可以根据自身项目特点和不同的发展阶段来选择适应的决策机制，在高效、透明、平权中选择实现某种平衡。因此，很难说哪一种决策机制更加先进，更加优秀。

贡献者指南：面向社区各类参与者有关代码、文档和其他贡献形式的，以及参与社区活动的综合性指南，其中不仅包括编码标准规范、代码审查流程和贡献工作流程，还包括参与社区建设和相关社区活动的行为准则，以及相应的权利受让和义务边界。

安全审查：涉及软件供应链安全管理、软件缺陷和漏洞管理，以及相应的机制和工具。

合规审查：涉及软件著作权、专利、商标和商业秘密相关的合规审查，其中重点是开源组件的开源许可证执行情况和这些开源许可证之间的兼容性，此外获得外部开发者的贡献者许可，和合作伙伴授予的商业许可的使用范围也是合规审查的重点内容。

开发者关系运营：（不仅是社区运营）包含了对核心开发者和外部开发者两类群体的关系运营。核心开发者包括基金会或企业内部存在雇佣关系的技术精英和少部分活跃的资深等社区开发者领袖，通常这类人群有较多的时间和精力投入到本社区，并具备较强的技术攻关能力和社区意见领袖这两个能力中的至少一种，他们是开源项目的主要贡献者和社区中流砥柱。满足核心开发者的个性化需求，为其创造更加便捷的工作条件是关系运营的重点。另一类外部开发者相较于前者，在研发能力和项目投入上均有不足，体现出来的特征为贡献少且不能持续，针对此类开发者关系运营的重点是为其降低社区进入的技术门槛和增加其内驱力，如增加更多的技术培训、提供更多的技术文档和技术支持，以及增加其未来在社区中的收获感会促使其更加持续的专注在这个项目上，各种精神奖励和物质馈赠，或者提供就业和创业机会都极具吸引力。

1.5.2 开源基金会的开源治理

开源基金会开源治理的核心目标是整合外部资源、参与技术创新、孵化所辖开源项目，进而推动开源生态的繁荣和发展。开源基金会通常是一个非营利性组织，一般由个人贡献者、企业赞助者和其

他利益代表组成的董事会来集体决策。

开源基金会具备公益属性，其开源治理往往更加强调开放、共享、透明和协作，更加注重开源布道，更加鼓励广泛的贡献者参与开源生态的共建。开源基金会常规开源治理流程如下：

1) 开源许可证知识产权的清查和管理

对开源项目进行许可证清查，列出使用到的所有许可证条款和限制条件，并进一步对这些许可证进行管理，确保该项目合法使用并遵守所涉及的所有开源许可证的要求，如GNU GPL、MIT License等。对捐赠方所取得的，与开源项目相关的知识产权进行审查。对开源项目的软件代码进行审查，确保未侵犯第三方的专利和商标等知识产权。

2) 开源项目的管理机构和管理流程制定

组建开源项目的管理委员会和技术监督委员会，明确开源项目的有关分发、贡献和管理的相关规定，并落实到管理委员会和技术监督委员会，确保捐献方、用户方、外部开发者和基金会内部都了解和遵守相关规定。

3) 定期进行开源项目供应链安全审查和合规管理监督

定期对开源项目所使用的第三方开源组件和库进行审查，确保其合法性和安全性，及时更新和修复可能存在的漏洞和安全问题，跟踪这些组件的来源、许可证和版本更新信息，确保所使用的组件符合合规要求。

4) 社区运营和项目推广

组建专业团队，调动开源基金会内部力量和外部资源运营开源项目，从人员、代码、社区三个维度，持续相应社区核心开发者、外部开发者群体、超级用户群体的需求，提高社区的活跃度和参与度，并积极主动推动该项目与供应链相关开源项目的合作和互动，促进技术交流和商业合作，以提升开源项目的可持续发展潜力。

5) 传播开源精神进行开源布道

定期面向社会各界传播开源文化，包括开源协作模式、主流开源技术、开源相关的风险意识等，主动传承开源文化，传播开源精神。

1.5.3 科技企业的开源治理

科技企业的开源治理服务于企业的整体战略，其特征为业务导向，即利用开源治理来服务于企业战略发展目标和商业利益。科技企业开源治理通常由企业管理层负责，决策权更为集中，但受到企业战略目标和商业利益的影响。

科技企业的开源治理往往侧重于企业面临的风险管控和开源项目的投资回报，以降低企业产品和服务的技术和合规风险，并增加企业竞争优势。科技企业常规开源治理流程。

1) 制定开源使用政策

确定企业如何使用和参与开源项目，包括如何选择使用开源项目、如何对开源项目进行贡献以及如何管理开源许可证。

2) 制定开源治理政策

建立企业内部的开源治理框架，明确决策制定的流程、相关参与部门的角色和责任。

3) 评估开源项目适用性

确定企业需求并选择适合的开源项目，评估项目的质量、社区活跃度、许可证兼容性等因素。

4) 进行风险评估

评估使用或贡献开源项目可能带来的技术风险、法律风险和供应链安全风险，并采取必要的措施进行风险管控。

5) 推动内部贡献

鼓励和支持企业员工参与相关开源项目，包括提供时间、资金奖励、资源和技术培训等。

6) 管理外部贡献

确保企业员工对外部相关开源项目的贡献符合公司政策和法律要求，可能包括代码审查、知识产权确权和许可证合规等。

7) 知识产权确权和许可证合规

通过技术手段持续监测相关开源项目的许可证信息和潜在的专利和商标隐患，明确相关开源项目的许可证权限，避免许可证冲突和违规使用，建立知识产权相关的主动防御机制和被动防范措施。

8) 进行安全漏洞管理

及时跟踪并处理开源项目中的安全漏洞，保护企业系统和数据的安全，避免给企业和用户方造成安全事故。

9) 建立与相关开源项目的合作关系

积极参与相关开源项目（供应链相关），建立良好的合作关系，促进企业在开源项目中的声誉和影响力。

10) 传播开源精神进行开源布道

定期对企业内部员工传播开源文化，包括开源协作模式和开源相关的风险意识等；定期举办和参与供应链相关的开源技术布道会，主动传承开源文化，传播开源精神。

第二章 开源开发者及项目发展现状

随着开源文化的普及和技术飞速发展，越来越多的开发者投身开源项目，共同促进开源生态的繁荣。作为项目核心动力，开发者从编写、测试到优化代码，持续推动项目进展。同时，他们的多元背景和专长，丰富了开源项目的技术内涵和功能应用。在生成式人工智能大趋势下，开发者积极吸纳新技术，显著提升了项目的智能化、性能、安全性和易用性，共同构建一个开放、协作、创新的生态体系。

近年来，中国开源生态发展迅速，呈现出以下主要特点：

开源人才大量供给：大量外企和国企的开源岗位，培养了一批开源人才，开源开发者增速很快。中国年轻一代的开发者具备全球化的视野和交际能力，开始主动融入全球协作，参与到开源项目。

开源社区发展迅速：涌现出一批具有影响力的开源社区，如开源中国、CSDN、掘金等，这些社区为开发者提供了丰富的技术资源、交流平台和互助环境，有力地推动了中国开源生态的发展。

开源项目数量逐年攀升：涉及云计算、大数据、人工智能、物联网等多个领域，部分项目已在国际开源社区中取得一定影响力。

2.1 开源开发者发展现状

2.1.1 开源开发者发展概述

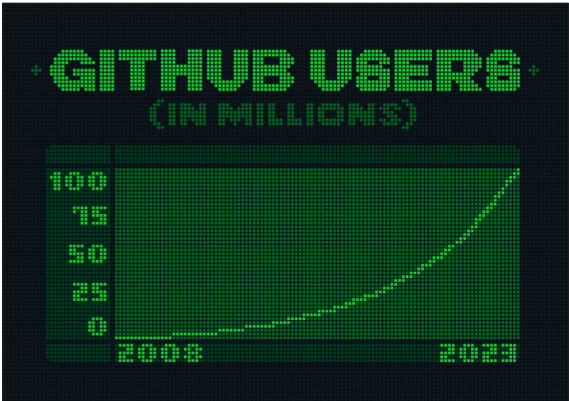
（1）开发者画像及规模、发展趋势

在2023年中国软件创新发展大会上，工信部发布了ICT产业的最新数据：中国的源代码贡献量已经达到世界第二，国内开源软件开发者数量突破800万，同样居全球第二。面对日益繁荣的开源市场，工信部已制定实施《“十四五”软件和信息技术服务业发展规划》，为加速开源发展和繁荣开源生态提出了明确战略部署和要求。此外，开源产业相关的中央及地方政策也日趋完善。

发布时间	发布单位	政策名称	关键内容
2023.12	江苏省人民政府	《关于加快工业软件自主创新的若干政策措施》	增强工业软件创新能力，支持开源技术创新。鼓励各地建设、招引自主开源软件创新平台，扶持基于自主基础软件的开源社区运营，举办开源技术交流活动，建立开源软件项目“白名单”制度，营造开源开放的技术产品创新和应用生态。
2023.10	南京市工业和信息化局	《加快开源软件发展三年行动计划（2023—2025年）》	组织实施开源软件发展“五大工程”，具体包括：开源项目发展工程、开源主体培育工程、开源应用牵引工程、开源公共服务能力建设工程、开源文化推广工程，构建开放共享、多方协同、融合创新、安全有序的南京开源生态。
2023.9	重庆市经济和信息化委员会	《重庆市加速培育软件开源创新生态助力中国软件名城建设实施方案》	培育优质开源项目，打造特色开源社区，加速集聚开源企业，完善开源治理生态，推进开源行业应用，以及加强对外开放合作。
2023.8	广州市人民政府	《广州市进一步促进软件和信息技术服务业高质量发展若干措施》	支持具有自主核心技术的开源社区、开源平台、开源项目发展，推动开源代码托管、代码检测等基础设施建设。加强与国际开源组织交流合作，推动商业开源软件研发，构建开源软件生态等。
2023.8	武汉东湖高新区	《加快促进软件和信息技术服务业创新发展的若干措施》	围绕开源服务平台、开源项目和开源协作针对性设计了支持政策；支持开源平台建设，单个平台项目最高可补贴3000万元；鼓励龙头企业发布自主技术开源项目，对原创性、爆炸性、轰动性的具有全球影响力的大模型，单个项目最高可补贴1000万元，形成大企业发榜、全球软件开发爱好者共同揭榜的氛围；同时着力提高开源社区活跃度，支持举办开源大讲堂、开源竞赛等活动。
2023.7	上海市人民政府	《立足数字经济新赛道推动数据要素产业创新发展行动方案（2023—2025年）》	推动功能型平台建设，打造大数据测试认证平台、关键共性技术综合测试平台，支持建设大数据开源平台和社区，构建数据要素专业服务体系。
2023.6	成都市经济和信息化局	《成都市关于进一步促进人工智能产业高质量发展的若干政策措施（征求意见稿）》	鼓励企业、科研机构围绕国产主流人工智能开发框架，研制AI芯片、开源函数库与工具包、中间件等，经评审择优，给予重点项目最高200万元经费支持。
2023.5	北京市经济和信息化局 北京市科学技术委员会 中关村科技园区管理委员会 北京市发展和改革委员会	《北京市通用人工智能产业创新伙伴计划》	要变革软件领域的开发应用模式，提升软件生产力工具效能，利用生成式AI重构企业软件，以开源聚合创新，构建大模型开源社区，吸引科研院所、代码托管平台、开发者及团队在京形成开放、包容、活跃的创新氛围。
2023.3	青岛市人民政府	《数字青岛2023年行动方案》	鼓励行业领军企业建设开源平台，孵化一批基础性、前瞻性开源项目，加快构建开源软件生态。加快推动卫星互联网等项目建设，探索研究卫星互联网与地面网络融合发展。
2023.3	济南市人民政府	《济南市加快软件名城提档升级三年行动计划（2023—2025年）》	支持企业积极参加国内开源基金会等开源组织增加研发投入，深度参与开源社区治理，鼓励企业及相关机构共同建设自主开源社区，探索完善各方联动的运营模式和可行的商业化路径。
2021.10	国务院	《“十四五”国家知识产权保护和运用规划》	健全大数据、人工智能、基因技术等新领域新业态知识产权保护制度。研究构建数据知识产权保护规则。完善开源知识产权和法律体系。
2021.3	国务院	《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》	支持数字技术开源社区等创新联合体发展，完善开源知识产权和法律体系，鼓励企业开放软件源代码、硬件设计和应用服务。

过去一年各地方政府的开源政策（图源：整理自网络）

中国的开源开发者数量呈现稳步上升状态。据GitHub 2023年数据显示，平台上已有1亿名开发者。根据GitHub Octoverse 2021报告显示，有755万用户来自于中国；在2022年新增的2050万用户中，近320万（约15.6%）来自中国。目前GitHub的1亿用户中，超过一千万名开发者来自中国。



GitHub平台开 发者增长趋势（图源：GitHub Octoverse）

从国家和地区分布来看，中国和印度的GitHub开发者用户数量和增长速度仅次于美国。中国开发者（含香港）在GitHub全球用户数量规模排名第三，2023年为1,100万，预计2026年GitHub中国开发者规模1,800 - 2,000万；但从活跃率（年度活跃数/总人数）可以看到，中国是前四名中最高的。

GitHub总用户量		美国	印度	中国（含香港）	中国大陆	中国香港	俄罗斯
1 32M		21.1M	14.2M	11.1M	9.3M	1.8M	3.2M
↑ 30%		↑ 22% 16.0%	↑ 34% 10.8%	↑ 11% 8.4%	↑ 7% 7.0%	↑ 34% 1.4%	↑ 21% 2.4%

全球开发者	年份	亚洲(M)	欧洲(M)	北美洲(M)	南美洲(M)	非洲(M)	大洋洲(M)
	2023	48.7 ↑ 27% 36.9%	42.4 ↑ 22% 32.1%	26.1 ↑ 21% 19.8%	8.0 ↑ 30% 19.8%	4.7 ↑ 34% 3.6%	1.8 ↑ 21% 1.4%
	2022	38.3 ↑ 28% 36.1%	34.7 ↑ 24% 32.7%	21.5 ↑ 23% 20.3%	6.2 ↑ 35% 5.8%	3.5 ↑ 40% 3.3%	1.5 ↑ 22% 1.4%
	2021	30.0 ↑ 30% 35.7%	27.9 ↑ 25% 33.2%	17.4 ↑ 23% 20.7%	4.6 ↑ 37% 5.5%	2.5 ↑ 40% 3.0%	1.2 ↑ 22% 1.4%
	2020	23.0 34.8%	22.4 33.9%	14.1 21.4%	3.3 5.0%	1.8 2.7%	1.0 1.5%

全球开发者国家/地区人数分布（Top15）（图源：GitHub官网）

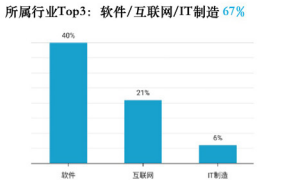
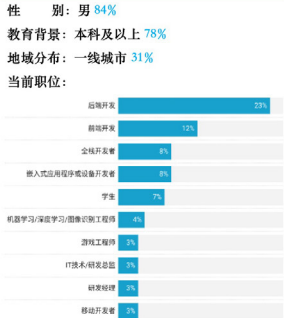
基于国内主流的技术社区及论坛的注册用户数，观察中国开发者的整体增长趋势时发现，中国开发者社区数据显示，中国开发者注册用户数4400万，2023年新增用户783万。随着开源的普及，很多年轻开源新生军逐渐加入和成长，开源人才增长空间和潜力巨大，覆盖全球95%的中文开发者，超过750万日均独立用户，超过1亿月独立访问用户。与此同时，CSDN于2022年推出面向新一代云原生开发者的开发云服务，为开发者提供更加稳定、便捷、高效的开源工具和开源服务。此外，CSDN携手华为云CodeArts推出新一代AI驱动的开源开发者平台，依托CodeArts一站式、全流程、安全可信的软件开发生产线，与CSDN数以千万的开发者，将成为人工智能驱动的新一代开发资源库和赋能平台，服务智能时代新开发者生态。目前用户数达到245万，项目超33万，入驻组织达到4.9万。

根据开源社区Gitee统计，2023年Gitee总用户数达到1200万，新增用户200万，总仓库数量达到3000万，新仓库数量500万。

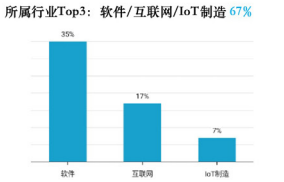
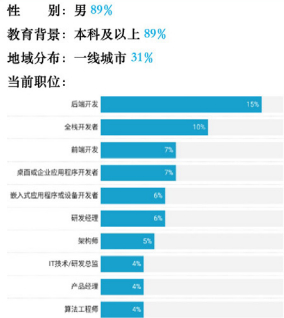
(2) 开发者行业、地域、技术特点等

根据CSDN发布的《2024中国开发者调查报告》显示，我们将开发者根据年龄范围划分成三类，对不同年龄段开发者群体进行特征分析后发现，30岁以下的开发者人数占72%，本科及以上学历占78%，84%的开发者都是男性。软件行业涵盖了国内四成开发者，40岁以下的开发者从事后端开发的比例最高。40岁以上开发者在一线城市占比最高，达到36%，所属行业重点分布在软件、教育和其他制造业领域。

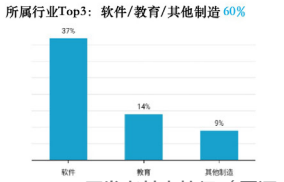
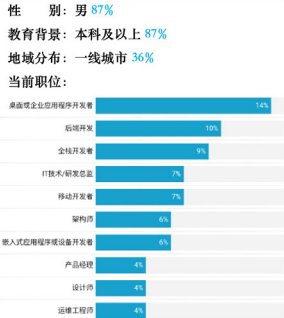
30岁以下的从业者（72%）



31-40岁从业者（22%）

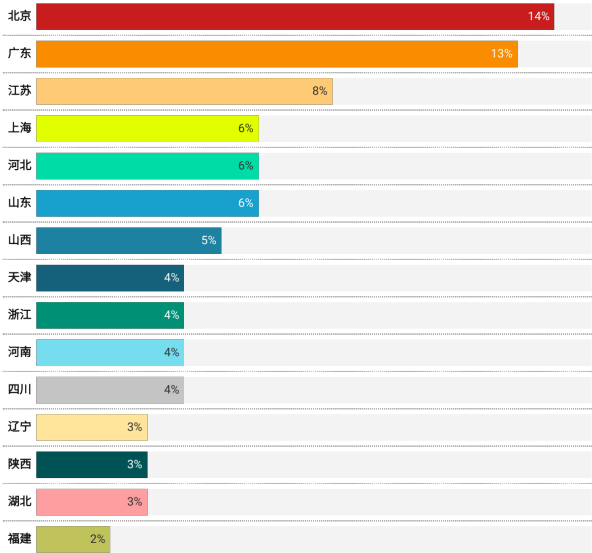


40岁以上的从业者（6%）



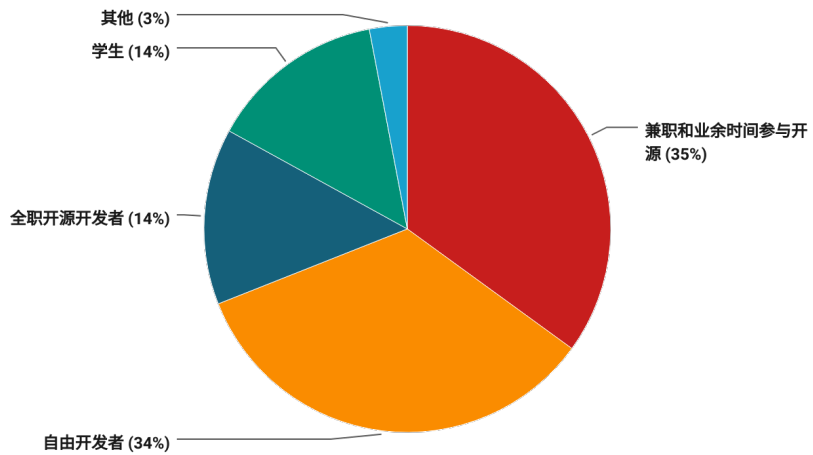
开发者基本特征（图源：CSDN）

全国超过40%的开发者工作在一线城市（其中以北京、上海、广州、深圳为主）。北京、广东是开发者聚集较多的地域，占全国总数的27%。江苏、上海地区的开发者占比数量处于第二梯队，占全国总数的14%。



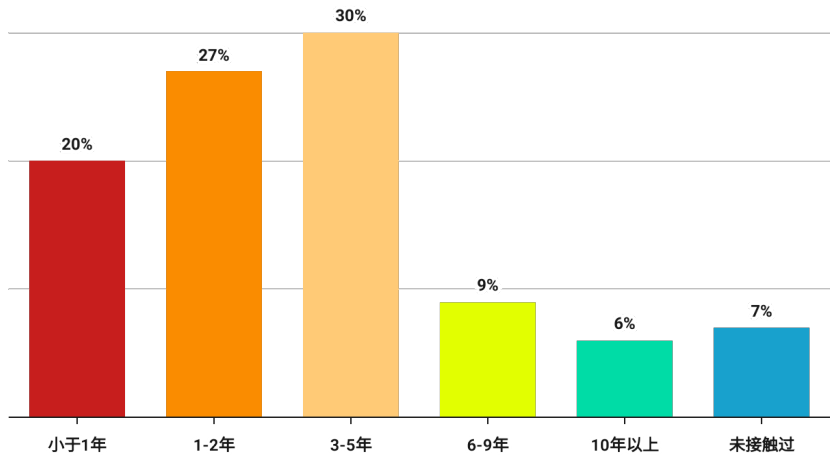
开发者地域分布（省、自治区、直辖市、特别行政区）Top15（图源：CSDN）

在参与开源项目的群体中，35%的人是兼职或在业余时间参与，34%是自由开发者，而全职开源开发者的比例为14%。这些数据反映了开源生态的多样性和广泛的参与基础，显示出即使在非全职的情况下，开源项目仍然吸引了大量贡献者，推动了技术的创新和社区的发展。



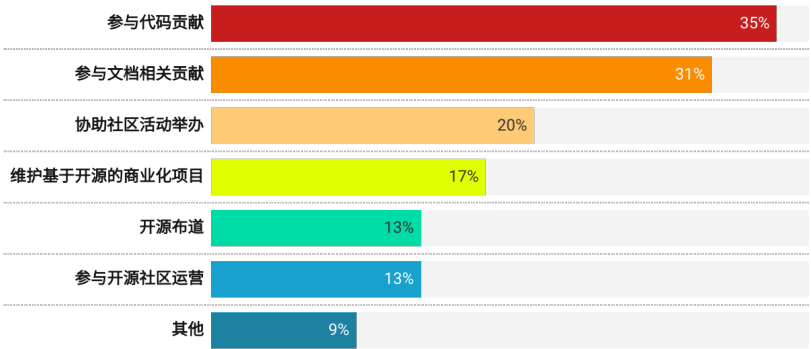
参与开源的人群比例（图源：CSDN）

越来越多的开发者通过参与开源项目提升技能、获取资源，并在全球开源社区中协作与分享。数据显示，近三成开发者接触开源已有3-5年的时间。



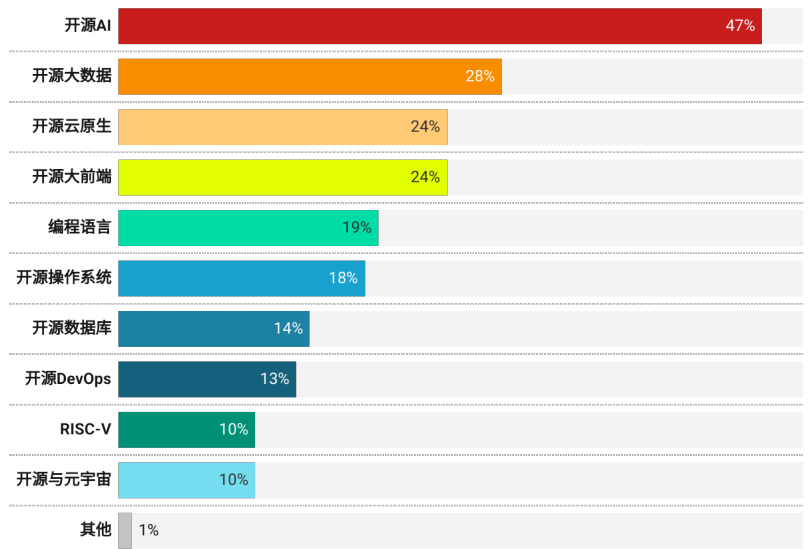
开发者参与开源的时间（图源：CSDN）

在开源项目的贡献上，主要还是通过代码和文档的贡献，在参与开源的开发者中，有35%的开发者有代码贡献，31%开发者有文档贡献。



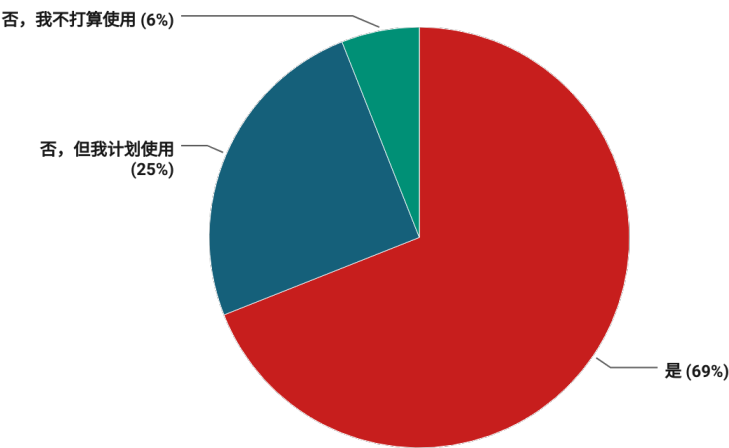
开发者参与开源的方式（图源：CSDN）

中国开源开发者在多个技术领域都有显著贡献。数据显示，开发者对开源技术领域的关注点主要集中在开源AI、开源大数据和开源云原生技术上，分别占比47%、28%和24%。这表明人工智能、数据处理和云计算是当前最受开发者重视的技术方向。总体来看，开发者对于新兴技术和基础架构的兴趣广泛，尤其是能够推动技术前沿发展的领域，如人工智能和大数据分析，显示了其在行业中的重要性和发展潜力。



开发者关注的开源技术领域（图源：CSDN）

全民AI时代，已有69%的开发者在日常工作、生活中用上了AI智能化工具。还有25%的人正处于计划使用中。



AI工具使用情况（图源：CSDN）

在全球开源背景下，超过75%的开发者支持AI大模型的开源，并认为这种开源模式可以促进全球开发者社区的参与和贡献，从而防止技术垄断。通过开源大模型，开发者能够共享资源和技术，实现创新的快速迭代。此外，开源大模型能够降低进入门槛，让更多中小企业和研究机构利用先进的AI技术，从而推动整个行业的进步和多样化发展。

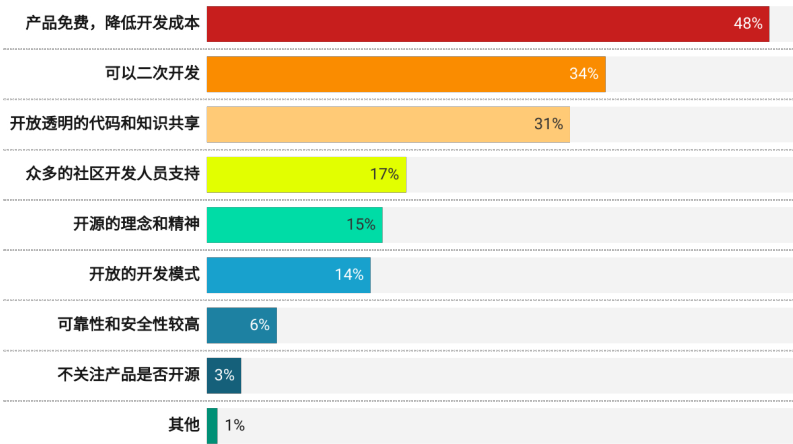
开源不仅有助于技术的普及，还能加强跨国界的合作，促进技术交流与共享。这种协作方式在提升项目的质量和影响力的同时，还为开发者提供了学习和成长的机会。

观点	占比
支持开源，开源大模型可以让全球开发者社区参与和贡献、防止垄断	75%
支持闭源，防范AI潜在风险，保持技术、商业领先	15%
不在乎开源还是闭源	10%

开源与闭源对AI大模型的影响（图源：CSDN）

2.1.2 开发者在全球开源的投入与贡献分析

选择开源的原因多种多样，其中48%的开发者认为开源协作方式能够显著降低开发成本，34%的人认为可以基于开源软件进行二次开发和定制自己的服务。这样的选择不仅节省资源，还能加速创新，促进技术交流和共享。在这种背景下，开源模式被越来越多的开发者和企业所接受和推广。除此之外，开放透明的代码和知识共享、众多的社区开发人员支持也是吸引开发者使用开源软件的重要原因。如果没有开源社区、开源活动、开源项目的支持，也不会有今天繁荣的开源生态。

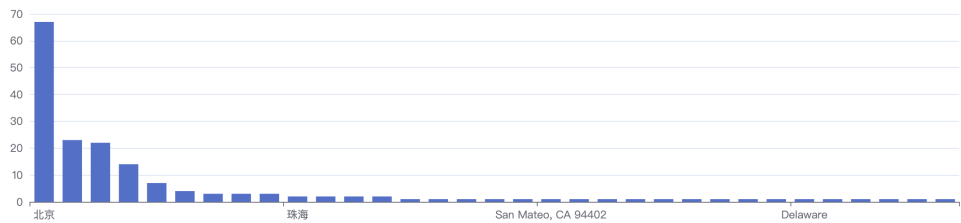


吸引开发者使用开源软件的因素（图源：CSDN）

(1) 开源组织和开源活动

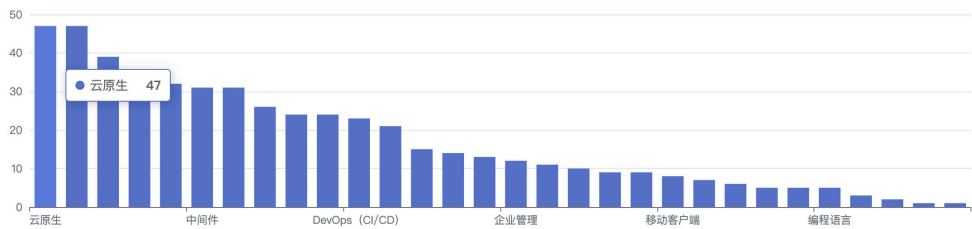
我国开发者发起和主导的开源社区不断发展，开源项目逐步实现国际化。数据显示，截至2022年，国内已经成立超过500个开源社区，涵盖了数据库、中间件、物联网、微服务、大数据、人工智能、移动互联网、云计算等技术领域。中国开发者主导的开源项目占全球12.5%，中国发起的开源项目海外贡献者占比33.7%。

就地域而言，多数社区集聚在北京地区，共有74个，相较去年新增了5个，代表社区有Apache IoTDB、Kube OVN、开源GitOps产业联盟、Casbin开源社区等。上海的开源社区数量排名第二，有27家，覆盖Kratos开源社区、Neo4j开发者社区、Graviti AI Community、OpenDataLab社区等。杭州地区有25家，包含Curve开源社区、Apache RocketMQ、Datawhale社区等。



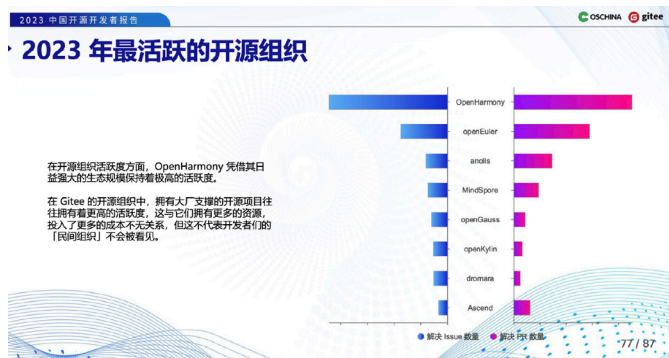
开源社区城市分布（图源：开源社）

今年云原生和大数据在开源社区领域的类别中数量最高，均为47个。这两种技术在当今数字化转型和企业运营中扮演着越来越重要的角色。云原生技术支持应用程序的可移植性和可扩展性，使企业能够更灵活地管理和部署应用。而大数据技术则是处理和分析海量数据的关键，帮助企业在数据驱动的决策中获得竞争优势。这些技术领域的项目数量之多反映了开源社区的高度活跃。开发者们在云原生和大数据方面的贡献显著增加，推动了技术的快速迭代和创新。开源社区的合作使得这些技术更加成熟和完善。



开源社区领域排行（图源：开源社）

在开源组织活跃度方面，据OSCHINA《2023中国开源开发者报告》显示，OpenHarmony凭借其日益强大的生态规模保持着极高的活跃度。其次分别为openEuler、anolis、MindSpore。



2023年最活跃的开源组织（图源：OSCHINA）

从2023年最受开发者关注的和喜爱的开源组织中，可以看出技术大厂凭借其技术底蕴和知名度仍然收获了大部分开发者的关注，但同时由开发者自发形成的开源组织受到了更多开发者的喜爱，如开源组织Dromara经过多年发展，已经逐渐有了自己的组织架构、项目捐赠机制和毕业机制，在用心的经营下，Dromara也在2023年共获得了超过57000个Star。



在过去一年，国内举办了多场开源社区活动和会议，为开发者提供了学习和交流的机会。例如：

全球开源技术峰会（Global Open-source Technology Conference），简称GOTC，是由上海浦东软件园、开放原子开源基金会、Linux基金会亚太区和开源中国联合发起的，面向全球开发者的一场盛大开源技术盛宴。GOTC 2023于上海举办为期2天的开源行业盛会，以行业展览、主题发言、特别论坛、分论坛的形式展现，与会者将一起探讨元宇宙、3D与游戏、eBPF、Web3.0、区块链等热门技术主题，以及开源社区、开源商业化、开源教育培训、云原生等热门话题，探讨开源未来，助力开源发展。

在2023年6月15~16日由COPU召开的《第18届开源中国开源世界高峰论坛》和《圆桌会议》上，中外专家对GPT的发展前景进行了热烈讨论。在会上谈到OpenAI、谷歌、微软等科技大公司在数据和模型权重方面拥有优势，结合他们的商业战略，选择了闭源大模型：OpenAI的ChatGPT，谷歌Bard、微软的AzureChatGPT，垄断了大型语言模型市场。座谈会结论性意见是：大模型走向开源的趋势几乎成为定局。

2023年6月11日-13日，2023开放原子全球开源峰会在北京圆满举办。本届峰会以“开源赋能，普惠未来”为主题，邀前沿科学家、技术领袖揭秘顶尖的开源趋势，与国际嘉宾、行业专家探讨开源领域最

前沿的思想与理论，与开发者大咖、技术创作者学习交流，聚集政、产、学、研、用、创、投、金等各领域的优势资源，共商开源发展大计，共筑开源发展未来，凝聚共建、共治、共享的发展共识。

2023年9月26日，由GOSIM开源创新汇主办，CSDN协办的GOSIM Conference在上海跨国采购会展中心隆重召开。来自英特尔、Arm、华为、微软、字节跳动、中兴、Futurewei、国汽智控、长安汽车、新加坡南洋理工、波特兰州立大学、清华大学、Rust社区等全球60+产学研专家助阵，进行技术解读和产业论证，旨在汇聚全球开源力量，提升开源价值，共铸开源新生态。

由开源社主办的第八届中国开源年会（COSCon'23）于2023年10月29日在成都圆满收官。这次大会众多开源大神云集，以“Mind Hackathon”为主题的创新主论坛形式让人耳目一新，同时分论坛延续往年传统将时下热点技术领域一网打尽，覆盖人工智能、开源商业、云计算、大数据、区块链、开源教育、开源文化、开源治理、开源硬件、开源操作系统、编程语言、女性论坛、Web应用开发、开源百宝箱以及闪电演讲等等。

2023 CCF中国开源大会于2023年10月21日至22日在湖南省长沙市北辰国际会议中心召开。本届大会的主题为“开源联合、聚力共赢”，邀请开源及相关领域中国科学院院士、中国工程院院士与知名专家，为大会带来特邀报告并参加开源高峰论坛等活动。本届大会作为中国开源领域的年度盛会，诚挚地邀请开源领域学术界、企业界、教育界的学者、从业者、师生等前来参会，共见中国开源新征程！

CSDN主办的第四届“长沙·中国1024程序员节”于2023年10月23-24日在长沙盛大举行，本届程序员节以“研发新高地，数智创未来”为主题，设置了岳麓对话、技术英雄会、十多场技术论坛以及编程大赛、创新展览、开发者嘉年华等精彩环节。议题涵盖人工智能、开源、AI模型、数据技术、AI编程与研发效能、自动驾驶与智能制造、信创等热点方向。

（2）开源项目开发者贡献发展现状

中国在开源项目上的发展近年来取得了显著成就。据X-Lab数据显示，全球开源仓库OpenRank排名如下：

注：OpenRank指标是由X-Lab开放实验室所研发，基于开源开发者一项目协作关系网络构建的协作网络指标，不仅能够很好地表征项目的整体发展状态、社区参与度，同时也引入了开源生态的要素，能够很好地将开源生态中位于关键协作位置的项目、人、组织等实体识别并展示出来。

🔍 全球开源仓库 OpenRank 排行榜 (Top 20)

#	仓库名称	OpenRank	参与人数	许可证	托管平台	国家
01	 NixOS/nixpkgs	22133.13  4139.52	6,915	MIT License	GitHub	
02	 home-assistant/core	15399.77  1290.72	18,057	Apache License 2.0	GitHub	
03	 microsoft/vscode	13167.44  1027.70	20,268	MIT License	GitHub	 United States of America
04	 MicrosoftDocs/azure-docs	10230.56  1135.47	10,291	Creative Commons Attribution 4.0 International	GitHub	
05	 pytorch/pytorch	10106.29  2080.33	6,697	Other	GitHub	 United States of America
06	 godotengine/godot	9742.78  2246.04	7,353	MIT License	GitHub	
07	 elidianaandrade/dio-lab-open-source	9552.18 	15,361	No License	GitHub	
08	 odoo/odoo	9233.43  2278.57	2,094	Other	GitHub	 Belgium
09	 flutter/flutter	9230.84  765.79	11,329	BSD 3-Clause "New" or "Revised" License	GitHub	 United States of America
10	 microsoft/winget-pkgs	8264.19  2519.74	2,261	MIT License	GitHub	 United States of America
11	 zephyrproject-rtos/zephyr	7939.92  1225.73	1,928	Apache License 2.0	GitHub	
12	 elastic/kibana	7561.98  283.28	1,427	Other	GitHub	 United States of America
13	 dotnet/runtime	7459.13  513.39	4,014	MIT License	GitHub	 United States of America
14	 mindspore/mindspore	7412.06  812.14	996	Apache License 2.0	Gitee	 China
15	 Expensify/App	7359.80  3830.93	1,921	MIT License	GitHub	
16	 firstcontributions/first-contributions	7214.80  1097.85	14,027	MIT License	GitHub	
17	 rust-lang/rust	7150.35  60.53	4,035	Other	GitHub	
18	 grafana/grafana	7037.03  875.99	4,513	GNU Affero General Public License v3.0	GitHub	
19	 llvm/llvm-project	6968.53  3699.19	4,582	Other	GitHub	
20	 vercel/next.js	6533.16  2415.76	10,872	MIT License	GitHub	

全球开源项目OpenRank排名 (Top20) (图源: X-Lab)

聚焦中国开源项目，OpenRank排名如下图所示。榜单上，位列前五的分别为openHarmony、openEuler、PaddlePaddle、MindSpore、Ant Design。这样的排名反映中国开源事业的蓬勃发展，企业与个人对于开源技术的贡献与参与度不断提升，同时也意味着开源项目在中国正受到越来越多的关注和实质性支持，无论是从政府、企业还是开发者层面。与此同时，在Top20榜单中，我们发现中国顶尖开源项目在操作系统、大数据、数据库、前端、人工智能等领域表现尤为突出。

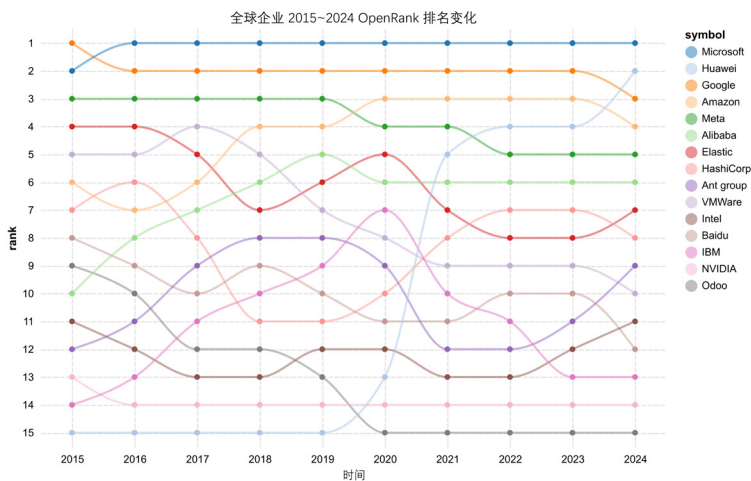
🔍 中国开源项目 OpenRank 排行榜 (Top 20)

#	项目名称	OpenRank	仓库个数	参与人数	托管平台	基金会
01	 openHarmony	34656.89  9430.84	741	4,949	Gitee	OpenAtom Foundation
02	 openEuler	18316.41  4253.45	10,164	3,945	Gitee	OpenAtom Foundation
03	 PaddlePaddle	15474.11  1078.97	121	7,893	GitHub & Gitee	
04	[M] MindSpore	10278.04  984.74	49	2,531	Gitee & GitHub	
05	 Ant Design	5377.15  326.47	133	6,159	GitHub & Gitee	
06	 TiDB	5138.51  605.96	32	768	GitHub	
07	 Apache Doris	4487.39  2076.74	8	1,125	GitHub	Apache Software Foundation
08	 openGauss	4418.85  1962.88	89	1,642	Gitee	
09	 Anolis	4346.94  1853.85	3,530	1,085	Gitee	OpenAtom Foundation
10	 StarRocks	3665.23  962.17	38	826	GitHub	
11	 Milvus	2986.18  594.79	62	1,232	GitHub	
12	 AntV	2845.46  29.00	117	2,150	GitHub	
13	 EMOX	2304.61  221.63	224	848	GitHub	
14	 TDengine	2067.48  11.61	38	923	GitHub	
15	 Apache Flink	2052.06  704.84	20	536	GitHub	Apache Software Foundation
16	 Apache ShardingSphere	1851.48  300.59	13	897	GitHub	Apache Software Foundation
17	 Apache Dubbo	1771.78  144.01	55	1,137	GitHub	Apache Software Foundation
18	 UmiJS	1637.41  474.79	73	1,906	GitHub & Gitee	
19	 MatrixOne	1479.43  564.20	36	96	GitHub	
20	 amis	1438.46  104.54	1	669	GitHub	

中国开源项目OpenRank排名 (Top20) (图源: X-Lab)

中国开源项目是以OpenDigger项目标签数据为准，单项目可能包含多个GitHub或Gitee平台上的组织或仓库。

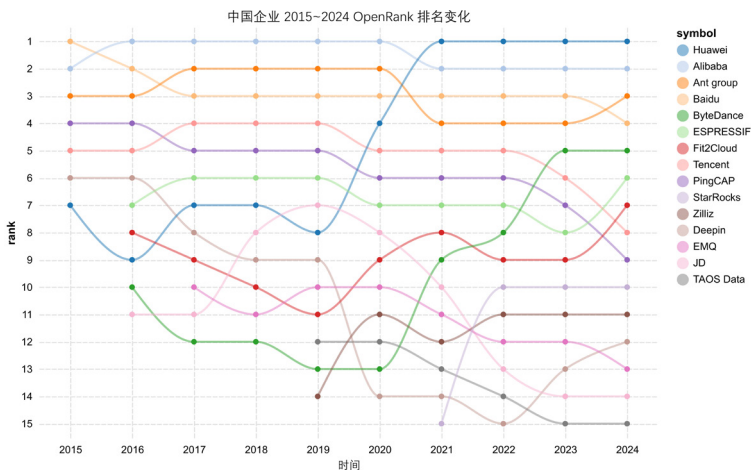
放眼全球重点参与开源的企业，排名前五的分别是微软、华为、谷歌、亚马逊和Meta公司。其中，微软自十多年前（2008年）开始布局开源，时任公司平台技术战略和开源软件实验室总监山姆·拉姆齐（Sam Ramji）在微软内部大力呼吁增强开源软件和闭源软件的互操作性。2014年，Satya Nadella担任微软首席执行官后，高调宣称：“微软爱Linux”，彻底改变了微软对开源的态度。在随后的两年多时间里，微软通过积极参与和贡献开源项目，于2016年达到了全球开源影响力的巅峰，并且这种影响力一直持续到今天，至今无人撼动。



全球企业OpenRank排名（Top20）（图源：开源社）

聚焦国内，华为于2019年将开源作为重要战略方向开始发力，仅用2年的时间就做到了国内第一，全球第二的位置；阿里巴巴在2021年前，一直是国内开源的领先者，并至今稳居世界第六的排名；蚂蚁集团在近三年的表现也是非常抢眼，并于2023年正式进入世界前十；国内开源的第四大巨头百度，由于国内快速的开源态势变化，目前排名全球12。

根据OpenLeaderboard榜单显示，进入全球前30的中国企业还有字节跳动（18）、PingCAP（19）、飞致云（24）、Deepin（25）、腾讯（26）、以及乐鑫科技（27）。



中国企业OpenRank排名变化（图源：开源社）

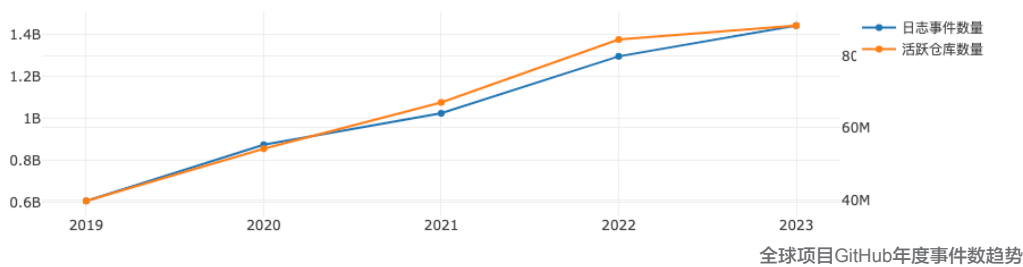
2.2 开源项目发展现状

2.2.1 中国开源项目概况

(1) 总体概况

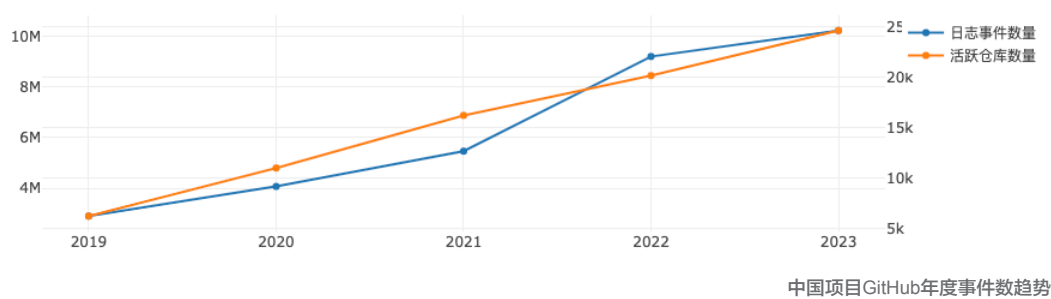
首先，统计分析全域GitHub的事件日志总体数量，如下图所示。

全球 GitHub 事件日志、活跃仓库五年趋势图



可以看到近几年全球开源的总体活跃情况和活跃仓库数量都在明显上升，显示了全球开源在快速发展。经过2018-2020年的高增长后，GitHub平台的年度事件增长数量逐渐下降。2023年GitHub日志数据为14亿，相比2022年增长了约10.32%。但由于整体体量的关系，10.23%的增长率还是一个非常大的数字，继续凸显出开源科技的发展在全球数字化转型中的活跃与关键地位。

中国项目 GitHub 事件日志、活跃仓库五年趋势图



在经历了疫情前后的快速增长后，GitHub平台上中国项目活跃度开始处于一个稳定的增长期。中国在2023年的日志事件和活跃仓库数均有明显的增长，相信在国家各项政策性意见的出台下，中国开源的发展会越来越好。

同时, 在全球的开源舞台上, 越来越多的中国代码托管平台涌现出来:

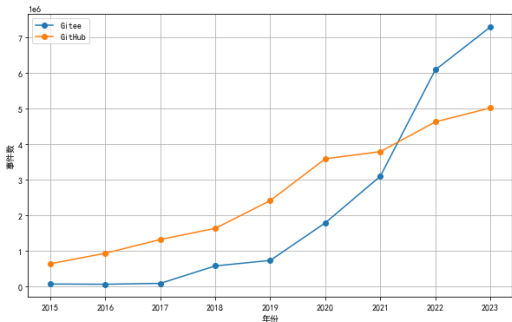
- **GitHub**: 拥有强大的协同功能, 能够让全球的开发者在同一项目中共同作业, 支持PR (Pull Request) 以及Issue等功能。同时, GitHub还提供了代码审查、项目管理、版本控制等一系列工具, 对于开源项目非常友好。其社区非常活跃, 有着丰富的第三方集成服务。然而, 对于私有项目, GitHub的价格相对较高。
- **GitLab**: 除了具有版本控制功能, 还具备项目管理、CI/CD等功能。GitLab有着完善的权限管理系统, 支持多人协同开发, 也可自建服务器, 适用于企业内部使用。其持续集成工具GitLab CI/CD功能十分强大, 可进行自动测试、部署等。不过, GitLab的界面相对于GitHub更复杂一些, 需要一些时间来适应。
- **Atomgit**: 2022年, 由开放原子开源基金会牵头组织阿里云计算有限公司、北京创新乐知网络技术有限公司、中国电子科技集团公司第十五研究所、北京航空航天大学等单位, 共同建设AtomGit平台。该平台旨在汇聚各方开源力量, 推动生态繁荣, 保障行业安全, 并长期专注于开源社区的建设。
- **GitCode**: 是CSDN携手华为云CodeArts推出新一代AI驱动的开源开发者平台, 依托CodeArts一站式、全流程, 安全可信的软件开发生产线与CSDN数以千万的开发者, 将成为人工智能驱动的新一代开发资源库和赋能平台, 服务智能时代新开发者生态。
- **Gitee**: 是开源中国社区2013年推出的基于Git的代码托管服务, 目前已经成为国内知名的代码托管平台, 致力于为国内开发者提供优质稳定的托管服务。除了提供最基础的Git代码托管之外, 还提供代码在线查看、历史版本查看、Fork、Pull Request、打包下载任意版本、Issue、Wiki、保护分支、代码质量检测、PaaS项目演示等方便管理、开发、协作、共享的功能。
- **GitLink**: 是CCF官方指定的开源创新服务平台, 旨在以“为开源创新服务”为使命, 以“成为开源创新的汇聚地”为愿景, 秉承“创新、开放、协作、共享”的价值观, 致力于为大规模开源开放协同创新助力赋能, 打造创新成果孵化和新工科人才培养的开源创新生态! 拥有分布式协作开发、一站式过程管理、高效流水线运维、多层次代码分析、多维度用户画像等多种功能。
- **Bitbucket**: 是Atlassian公司开发的一款Git版本控制系统, 与Jira、Confluence等工具有很好的集成, 方便团队协同开发。Bitbucket支持私有仓库, 针对小团队有免费方案, 比较适合初创公司使用。然而, 相较于GitHub和GitLab, Bitbucket的社区活跃度较低。
- **SourceForge**: 这是一个开源软件开发和下载平台, 允许用户利用Git进行版本控制, 并提供项目管理、错误跟踪、功能请求等工具。其界面设计较为陈旧, 开源项目的数量和活跃度相较于GitHub要低

一些。

- **Beanstalk:** 一款为企业设计的Git项目管理工具，拥有代码审查、发布工具、时间线、通知和报告等功能，还提供了强大的权限管理功能和集成其他服务的能力，但价格较高，不太适合小团队或者个人开发者。
- **Gitea:** 一个开源的自托管Git服务，界面简洁，部署简单，支持多人协同开发，拥有issue跟踪和管理系统，更适用于小型团队或者个人开发者。
- **Coding:** 是一个面向开发者的云端开发平台，致力于做项目众包平台。支持多种账号登陆，实行用户等级制。
- **阿里云效Codeup:** 阿里云的代码管理平台，致力于保障企业代码资产安全，支持自动化代码扫描、合并评审管控等高级功能。
- **腾讯Coding:** 腾讯旗下的开发平台，提供一站式开发管理协作工具，帮助研发团队快速落地敏捷开发与DevOps开发方式，实现研发效能升级，适用于多种研发工作场景。
- **SVNChina:** 国内成立较早的支持私有项目的SVN源代码托管平台。
- **Google Code:** 拥有良好的issue tracker、wiki系统以及Google group/gmail的集成，同时支持svn/mercurial/git三种版本管理，但在中国大陆无法访问。

(2) GitHub和Gitee的总事件数趋势比较

由于GitHub平台活跃事件的庞大，接下来的分析工作，建立在每个平台的前3万个活跃仓库的基准之上。为了方便比较，选取了GitHub与Gitee中与开源参与相关性较大的8类事件进行统计分析，包括CommitCommentEvent、ForkEvent、IssueCommentEvent、IssuesEvent、PullRequestEvent、PullRequestReviewCommentEvent、PushEvent和WatchEvent。

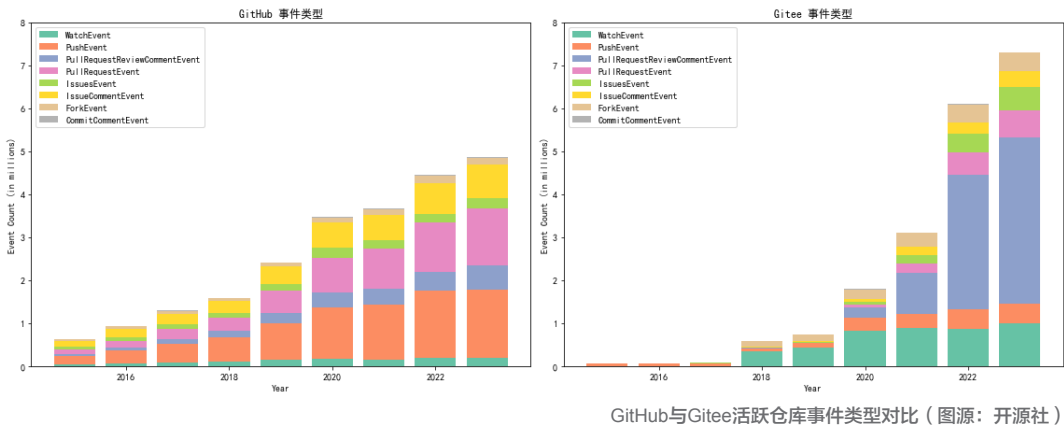


GitHub与Gitee活跃仓库事件数（图源：开源社）

分析发现Gitee平台呈现更为显著的增长趋势。甚至从2021年起，前3万个活跃仓库的事件数量超越了GitHub，凸显了国内活跃开源项目的爆发态势。反映了国内开发者积极参与和贡献到开源社区的热情，为技术创新和知识共享注入了新的活力。

然而，需要强调的是，单纯依靠前3万个活跃项目的数据无法完全揭示全球GitHub平台的真实情况，因为长尾效应在全球范围内仍然非常明显。特别是GitHub平台作为全球领先的开源社区的广泛和多样性。在未来，随着技术的不断演进和开源文化的推动，可以期待中国开源社区在全球范围内继续蓬勃发展。

再进一步，分析基础事件的细分领域数据，结果如下图所示。



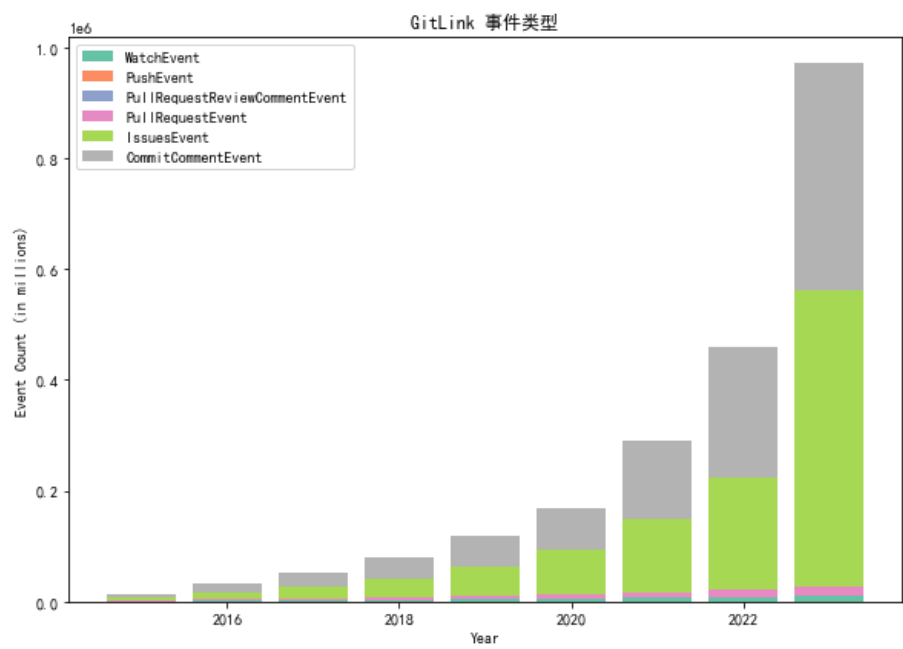
从分析结果可以看到：

在GitHub平台上，最多的事件类型是Push事件，Pull Request事件和Issue Comment事件分别居二三席。其中各个事件的发生次数占比基本上没有太大的变化，这体现出GitHub的开源生态模式在走向一个稳定的趋势。

在Gitee平台上，事件数据在2018年有极大的增长，最初以Watch事件为主。但在2020年后，Pull Request、Review Comment事件开始快速增长，在2022年成为最多的事件类型，并且在2023年也有持续的增长。Gitee事件数据的结构性变化，体现出国内开发者从关注者到贡献者角色的巨大转变，这和全球范围内的观察都是一致的。

(3) GitLink事件数分析

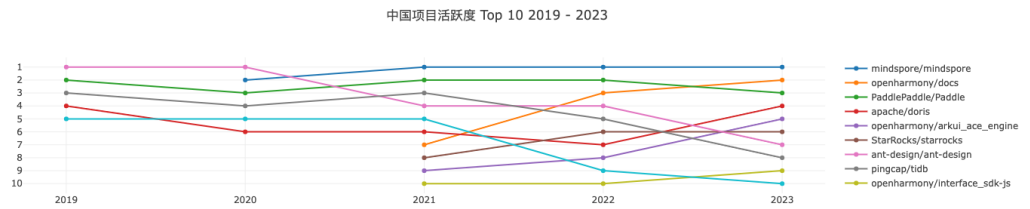
对于GitLink平台，同样选择了前3万个活跃仓库作为基准。鉴于数据的局限性，仅选取了包括CommitCommentEvent、ForkEvent、IssueCommentEvent、IssuesEvent、PullRequestEvent和WatchEvent六种事件类型的数据进行分析。



GitLink平台事件数据分析（图源：开源社）

观察到，尽管Gitlink的活跃仓库事件数相较GitHub、Gitee等平台仍存在一定差距，但同样呈现明显的增长趋势。在Gitlink平台上，Issue事件和CommitComment事件占活跃仓库事件的绝大多数。

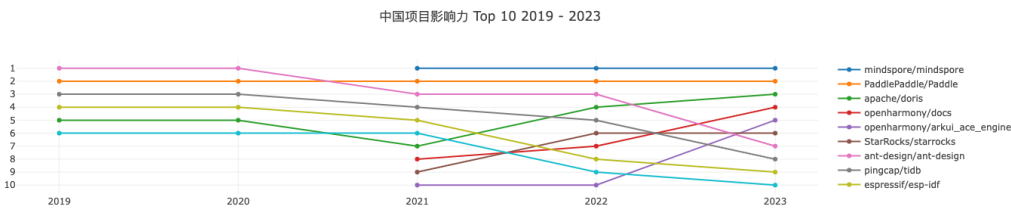
2.2.2 中国项目活跃度与影响力趋势分析



中国项目活跃度Top10 2019-2023

可以看出，从2021年开始MindSpore项目上升至第1。MindSpore是一种适用于端边云场景的新型开源深度学习训练/推理框架。

OpenHarmony活跃度上升至第2，StarRocks得益于其良好的开源生态以及全球的推广，在2021年上榜，并在2023年直接上升至第6位。



中国项目影响力Top10 2019-2023

MindSpore稳居中国开源项目影响力第一，Paddle长期居于第二，OpenHarmony影响力上涨较快，Doris项目近3年排名上涨较快，反映出中国开源项目积极的发展。

2.2.3 中国项目活跃度与影响力榜单

在开源项目洞察这部分，同样根据活跃度和影响力，分别比较了全球和中国排名前20的开源项目，可以看到：

在单个项目上，中国主导的开源项目的活跃度和影响力距离全球顶级开源项目的活跃度与影响力还有一定的差距，特别是在影响力方面，需要长期建设，需要最大限度融入全球开源生态里。

中国主导的开源项目中，华为旗下的相关项目排名相对靠前，mindspore/mindspore活跃度与影响力均位列第一。百度旗下PaddlePaddle相关项目也表现不俗，pingCAP旗下的TiDB与StarRocks旗下的StarRocks数据库排名相对靠前。

排名	项目	活跃度	托管平台
1	ghscr/ghscription	86763.63	GitHub
2	NixOS/nixpkgs	56623.46	GitHub
3	home-assistant/core	44216	GitHub
4	digitalinnovationone/dio-lab-open-source	43453.41	GitHub
5	microsoft/vscode	39513.63	GitHub
6	firstcontributions/first-contributions	38068.69	GitHub
7	MicrosoftDocs/azure-docs	30196.24	GitHub
8	flutter/flutter	29889.87	GitHub

9	AleoHQ/leo	29009.35	GitHub
10	pytorch/pytorch	27393.73	GitHub
11	godotengine/godot	25421.53	GitHub
12	llvm/llvm-project	24764.32	GitHub
13	ibm-developer-skills-network/jbbmo-Introduction-to-Git-and-GitHub	22857.21	GitHub
14	langchain-ai/langchain	22748.28	GitHub
15	microsoft/winget-pkgs	21831.91	GitHub
16	Expensify/App	21548.12	GitHub
17	odoo/odoo	20748.32	GitHub
18	dotnet/runtime	20710.31	GitHub
19	vercel/next.js	20662.23	GitHub
20	elastic/kibana	20546.96	GitHub

表1 全球项目活跃度Top20

排名	项目	活跃度	托管平台
1	mindspore/mindspore	16028.37	Gitee
2	openharmony/docs	15969.05	Gitee
3	PaddlePaddle/Paddle	13865.37	GitHub
4	apache/doris	12290.44	GitHub
5	openharmony/arkui_ace_engine	11985.97	Gitee
6	StarRocks/starrocks	10508.79	GitHub
7	ant-design/ant-design	8978.52	GitHub
8	pingcap/tidb	8036.56	GitHub
9	openharmony/interface_sdk-js	6854.03	Gitee
10	milvus-io/milvus	6520.56	GitHub
11	espressif/esp-idf	6326.43	GitHub
12	openeuler/docs	6000.69	Gitee
13	openharmony/xts_acts	5355.52	Gitee
14	PaddlePaddle/PaddleNLP	4722.73	GitHub
15	PaddlePaddle/PaddleOCR	4690.18	GitHub
16	openharmony/graphic_graphic_2d	4606.45	Gitee
17	apache/flink	4412.76	GitHub
18	taosdata/TDengine	4320.91	GitHub
19	apache/shardingsphere	4150.35	GitHub
20	matrixorigin/matrixone	4031.98	GitHub

表2 中国项目活跃度Top20

排名	项目	活跃度	托管平台
1	NixOS/nixpkgs	22133.13	GitHub
2	home-assistant/core	15399.77	GitHub
3	microsoft/vscode	13167.44	GitHub
4	MicrosoftDocs/azure-docs	10230.56	GitHub
5	pytorch/pytorch	10106.29	GitHub
6	godotengine/godot	9742.78	GitHub
7	elidianaandrade/dio-lab-open-	9552.18	GitHub
8	odoo/odoo	9233.43	GitHub
9	flutter/flutter	9230.84	GitHub
10	microsoft/winget-pkgs	8264.19	GitHub
11	zephyrproject-rtos/zephyr	7939.92	GitHub
12	elastic/kibana	7561.98	GitHub
13	dotnet/runtime	7459.13	GitHub
14	mindspore/mindspore	7412.06	Gitee
15	Expensify/App	7359.8	GitHub
16	firstcontributions/first-	7214.8	GitHub
17	rust-lang/rust	7150.35	GitHub
18	grafana/grafana	7037.03	GitHub
19	llvm/llvm-project	6968.53	GitHub
20	vercel/next.js	6533.16	GitHub

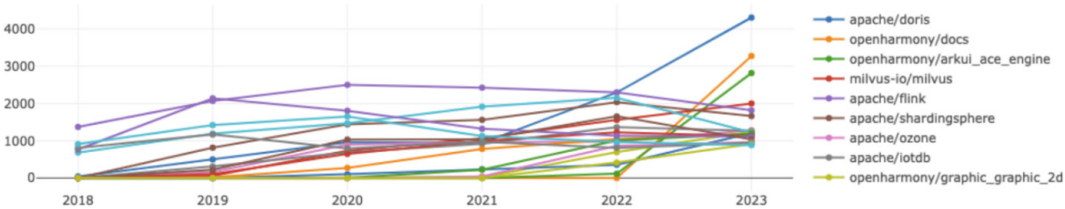
表3 全球项目影响力Top20

排名	项目	活跃度	托管平台
1	mindspore/mindspore	7412.06	Gitee
2	PaddlePaddle/Paddle	5393.8	GitHub
3	apache/doris	4262.44	GitHub
4	openharmony/docs	3990.5	Gitee
5	openharmony/arkui_ace_engine	3328.36	Gitee
6	StarRocks/starrocks	3167.61	GitHub
7	ant-design/ant-design	3022.38	GitHub
8	pingcap/tidb	2186.94	GitHub
9	espressif/esp-idf	2132.46	GitHub
10	milvus-io/milvus	1973.63	GitHub

11	apache/flink	1787.59	GitHub
12	taosdata/TDEngine	1688.05	GitHub
13	apache/shardingsphere	1661.84	GitHub
14	PaddlePaddle/PaddleNLP	1653.6	GitHub
15	PaddlePaddle/PaddleOCR	1471.01	GitHub
16	baidu/amis	1438.46	GitHub
17	openharmony/xts_acts	1381.53	Gitee
18	openharmony/graphic_graphic_2d	1367.68	Gitee
19	openeuler/docs	1365.86	Gitee
20	metersphere/metersphere	1335.11	GitHub

表4 中国项目影响力Top20

2.2.4 基金会旗下中国项目影响力趋势分析

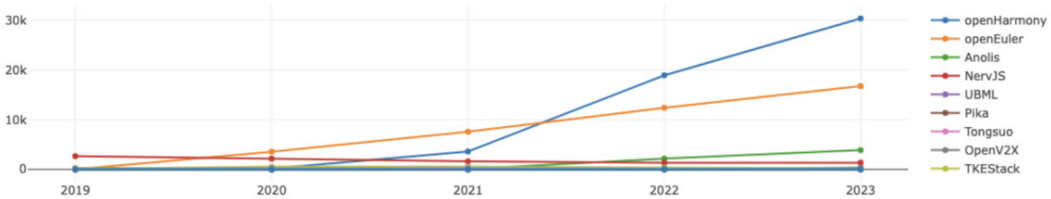


中国基金会项目OpenRank变化趋势

各基金会旗下的中国项目单独分析来看：

- Doris和OpenHarmony发展速度最为抢眼；
- Milvus向量数据库由于AIGC领域的需求，发展迅速；
- Flink、ShardingSphere等项目稳中有降。

开放原子开源基金会旗下项目OpenRank趋势分析：



开放原子开源基金会项目OpenRank变化趋势

今年能够第一次看到开放原子旗下各项目的发展情况:

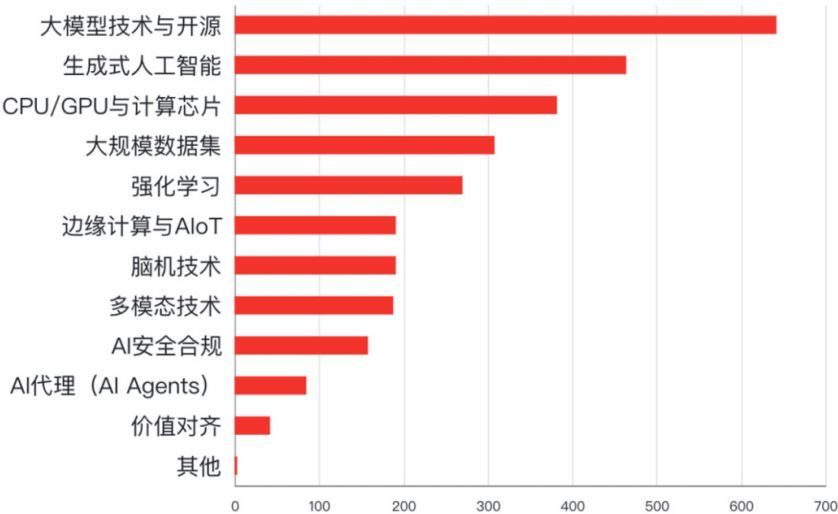
- 排名前三的分别是OpenHarmony、openEuler和Anolis，代表了操作系统绝对的地位，特别是OpenHarmony，发展最为迅速；
- 其他上榜项目发展非常平稳，期待后续能够在新的一年里中发力。

2.3 开发者面临的挑战、建议与未来展望

2.3.1 面临的挑战

随着开源软件的蓬勃发展，以及人工智能技术的逐渐成熟，也为中国开源领域带来了巨大的挑战。

其一，大模型技术、开源、生成式AI、算力这些热词都是开发者最关注的技术突破方向。而更深远的问题，比如与隐私保护、日常生活息息相关的价值对齐、安全合规等问题，受关注度弱于其他技术突破方向。这可能是因为很多人还未对「人工智能的潜在伦理风险」产生实感，这些情节暂时还存在于电影之中。

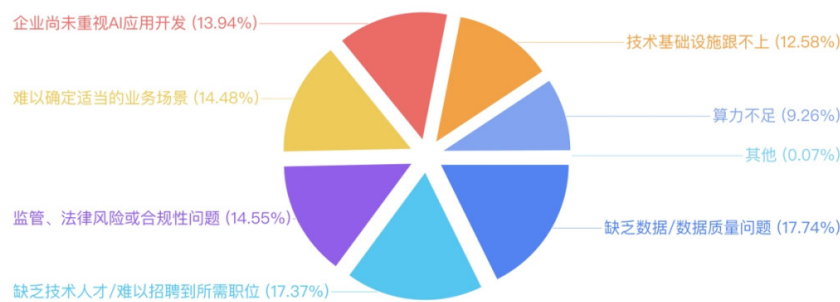


开源开发者关注的人工智能突破方向（图源：CSDN）

隐私保护也逐渐成为一个不可忽视的挑战。目前，以大模型产品为例，国内外诸多公司也遇到过因“用公开数据训练大模型”导致官司缠身，这也引发了巨大的争议。开源AI领域需要处理大量的数

据，但同时也需要保护数据的隐私和安全性。随着对个人数据使用的监管加强，保护用户隐私成为一个重要挑战。

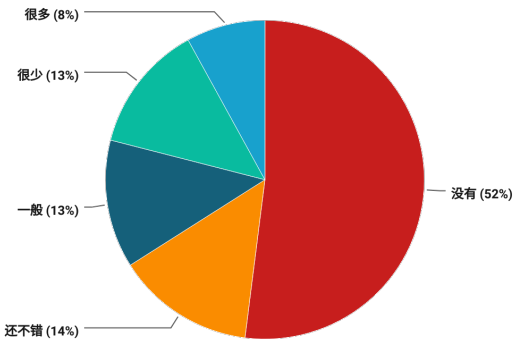
根据CSDN发布的《AI开发者生态报告》显示，在大数据时代，尽管数据量呈指数级增长，但获取高质量的数据、处理数据偏见、保障数据安全与隐私，以及有效利用数据进行创新研究等方面依然面临巨大挑战。



开发者眼中的发展挑战（图源：CSDN）

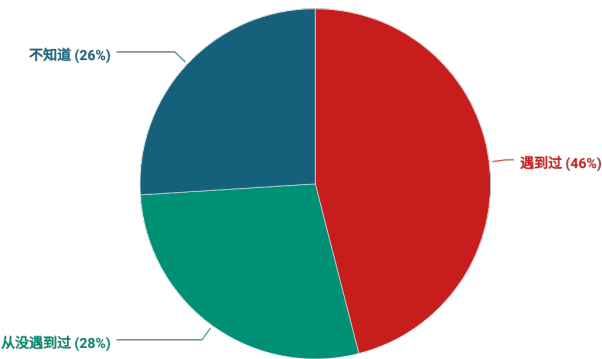
其二，知识产权和许可管理的挑战，导致代码滥用或滥用风险。开源软件项目需要清晰的许可协议，但有时会存在许可混乱或侵权问题，这引发法律纠纷和不确定性。 在过去一年里，也迎来了一则好消息。2024年2月，最高人民法院发布了一则《这份判决给软件开发者吃了定心丸》文章，最高法知识产权法庭平衡保护软件开源社区建设和软件开发者权益受好评。“最高人民法院对案件的妥善处理，合理地界定了GPLv2（开源协议）此类开源许可证软件二次开发者的权利边界，为我国正在建设有中国特色的开源社区树立了标杆，是我国软件开发行业的重要里程碑，意义非凡。”这是良好的开端。

其三，维护和支持的可持续性，一些项目缺乏长期支持和资金，让很多开源开发者面临生存问题，最终导致一些项目难以长期维护，会停滞或废弃，影响用户和企业的可靠性。在这一窘境下，52%的开发者称自己没有



开发者在开源上获得的收入（图源：CSDN）

其四，安全性和漏洞管理。开源软件虽然公开代码，但并不保证免受漏洞和安全威胁的影响。不及时发现和修复漏洞会给用户和组织带来风险，特别是在广泛使用的开源项目中。调查数据显示，46%的开发者遇到过开源软件相关的安全漏洞。



开源软件的安全漏洞（图源：CSDN）

2.3.2 建议

基于以上挑战，开源开发需要包括政府、企业、社会组织以及个人开发者的积极参与和支持，以下是一些建议和措施：

(1) 加强数据隐私与安全性

- 数据最小化原则:在项目中最小化收集和使用个人数据。只收集和使用必要的的数据，避免收集不必要的个人信息。
- 用户明示同意: 在收集个人数据之前，确保获得用户的明示同意。用户应该清楚地知道他们的数据将被用于何种目的，并且有权选择是否同意。
- 匿名化和脱敏: 在可能的情况下，对收集的个人数据进行匿名化或脱敏处理，以降低数据关联性和识别性。
- 数据访问权限控制: 限制对用户数据的访问权限，仅允许有必要权限的人员或系统进行访问，并建立审计机制以监控数据访问情况。

- 社区教育和意识提升：加强社区成员对隐私保护的意识，教育他们如何保护个人隐私，并提供相应的培训和资源。

(2) 采用有效的方式更好地管理知识产权和许可证，降低法律风险

- 开源项目应该制定明确的知识产权政策，包括许可证选择、版权声明、贡献者权利和责任等方面的规定。这有助于消除不确定性，增加项目的透明度和可预测性。
- 选择合适的许可证：选择一个适合项目需求和目标的开源许可证是至关重要的。开源社区提供了多种选择，如MIT、GPL、Apache License等，每种许可证都有其特定的特点和限制。项目组织者应该认真评估每种许可证的优缺点，选择最适合项目的许可证。
- 明确贡献者的权利和责任：确保贡献者清楚地了解他们的贡献如何受到保护，并明确他们对项目的贡献的版权归属。这可以通过贡献者协议或贡献者许可协议来实现。
- 建立贡献者指南和流程：提供清晰的贡献者指南和流程，包括如何提交贡献、如何报告问题、如何进行代码审查等。这有助于规范贡献的方式，减少知识产权纠纷的可能性。
- 定期审查和更新许可证：定期审查项目使用的开源许可证，并根据项目的发展和需求进行更新。随着时间的推移，项目的需求和环境可能会发生变化，因此需要不断地审查和更新许可证选择。
- 采用自动化工具：使用自动化工具来管理知识产权和许可证信息，如软件扫描工具、许可证管理工具等。这些工具可以帮助识别项目中使用的开源组件和其相关的许可证信息，从而更好地管理知识产权和遵守许可证要求。

(3) 维护和支持的可持续性

- 多样化的资金来源：探索多样化的资金来源，包括个人捐赠、企业赞助、基金会资助等。建立稳定的资金来源有助于项目的长期发展和运营。
- 建立商业模式：考虑建立基于项目的商业模式，例如提供付费支持和咨询服务、开展培训和认证等。通过商业模式来获得收入，可以支持项目的持续发展和维护。
- 定期更新和维护：定期更新和维护项目，修复漏洞、改进功能、更新文档等。这有助于保持项目的活跃度和吸引力，增加用户和贡献者的参与度。

- 建立社区参与机制：建立开放、包容的社区参与机制，鼓励更多的人参与项目的开发和维护。这不仅能够分担项目维护的压力，还能够增加项目的可持续性。

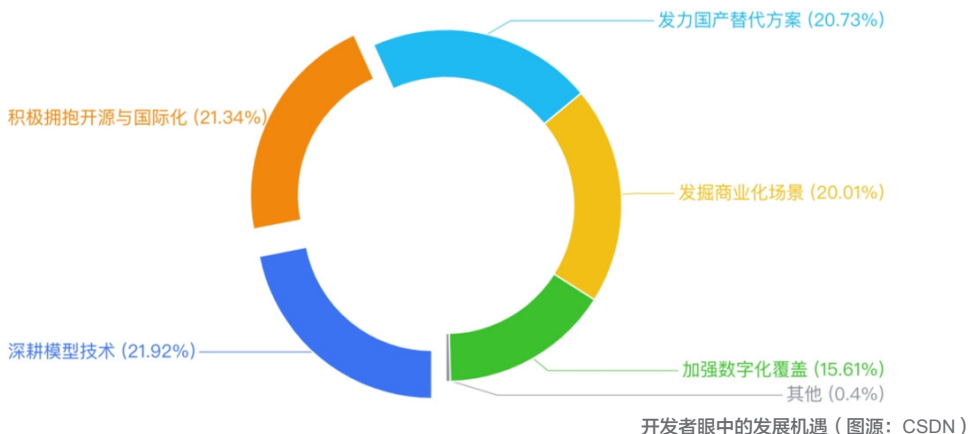
（4）使用内存安全编程语言

随着技术的进步和需求的变化，开发人员逐渐意识到选择适合的编程语言来减少内存管理的风险是明智之举。使用那些在设计时就注重内存安全的语言，可以通过自动内存管理、严格的类型检查和内置的安全特性，显著降低了开发人员引入内存错误的风险。

例如，Rust通过借用检查器和所有权系统，确保在编译时捕获几乎所有的内存错误；Go语言通过垃圾回收机制有效地管理内存；Java同样利用垃圾回收，确保对象的生命周期管理更加安全；Python和Ruby则通过高层次的抽象和动态类型系统，减少了内存管理的复杂度。Swift不仅为iOS和macOS开发提供了强大的功能，同时也通过自动引用计数（ARC）来实现内存安全。

2.3.3 未来展望

随着数字化转型的不断推进，越来越多的企业和组织意识到开源软件的价值，其在削减成本、提高安全性、推动创新和实现可持续性方面的优势日益显著。开源开发者将继续在推动开源和开源人工智能领域发展中发挥重要作用。他们将继续引领创新，推动技术进步，并积极参与社区建设，促进合作和知识共享。



在AI维度，开发者眼中的发展机遇同样丰富多彩，拥抱开源、出海与国际化、深耕模型技术、发力国产替代、挖掘商业场景、加强数字化建设……这些都是亟待积极把握的多元化发展机遇。

第三章 开源重点领域发展现状

3.1 操作系统产业研究报告

3.1.1 调研背景

操作系统是信息化的核心基础设施，自1991年Linux操作系统发布以来，开源一直是我国发展操作系统产业的重要路径。

经过多年的探索和尝试，第一代Linux中文发行版开始陆续推向市场，与上游社区的割裂成为这一带产品的典型特征也为后续的自主发展和演进埋下了隐患。后续中国操作系统企业的发展是相对艰难的，热情冷却后要面临的是自我的救赎，这时我们大多数人开始意识到“开源并不简单”。2008年是中国奥运之年，是中国向世界打开的一扇窗，这一年也是中国开源和中国操作系统的转折之年，在中国开源泰斗陆首群主席的努力斡旋下，Linux基金会首次在华召开了2008 Linux Symposium，同年Apache软件基金会、openOffice基金会、Gnome基金会也在华完成了它们的亚洲首秀，中国开源开发者和中国Linux企业在全球开源界的形象大幅提升，为后续进一步融入全球开源大家庭奠定了基础。

2019年3月，正当我们为中国开发者内核贡献度全球排名连续排名第一（自2017年11月Linux-4.14发布持续至今）而欣喜的时候，美国对中国开始了史无前例的科技封锁和技术供应链解耦，以华为为首的中国科技企业面临有史以来最大的生存危机。如操作系统等信息产业核心技术在民间瞬间被“卡脖子技术”一词代替。所幸的是，我国几十年来在开源生态中积累了大量的人才和工程经验，同时信创市场给国产软硬件企业撑起了一个稳定的市场空间，保证部分国内企业可以不断发展，不断迭代技术产品。

今天，华为、阿里、腾讯、麒麟、统信、睿赛德、翼辉信息等企业的相关操作系统产品的用户量持续增长，品牌影响力大幅提升。2022年受COPU委托，北京初心开源技术有限公司面向整个行业开展了一次中国操作系统产品调研。2023年10月至2024年3月，在第一次调研数据的基础上，COPU联合北京开源创新委员会、北京初心开源技术有限公司、清华大学计算机系联合发起范围更广的中国操作系统调研。

本次调研希望能够完整呈现当前我国操作系统产业发展的整体情况，产品分布，技术特点和现存问题。

3.1.2 调研方法

本次中国操作系统产品调研分为两个阶段，第一个阶段面向全社会广泛征集现已对外发行的操作系统产品（包含同一单位的多个发行版，不包含内部使用不对外发行的自用操作系统），并对其产品特性进行标注和补充。

第二个阶段是定向邀请来自30家不同企业和高校的操作系统专家对第一阶段征集的操作系统产品进行初选,再通过技术栈成熟度和发展性、系统和应用生态开放性、存储性能、网络性能、硬件性能等十五个维度进行系统考量。

3.1.3 调研数据

本次调研数据《中国操作系统产品名录》为公开文档(如下图)。

[illegible]

截止2024年3月30日，本次操作系统调研（去重后）共收录各类中国操作系统发行版累计216个，其中桌面类24个，手机平板类39个，云和服务类38个，物联网类71个，汽车类28个，嵌入式类46个，特种应用类15个。

桌面类	手机平板类	云和服务类	物联网类	汽车类	嵌入式类	特种类
24	39	38	71	28	46	15
注：上述分类有交叉，计数有重复						

相较于2022年9月第一次调研数据，中国操作系统发行版数量增长18.68%，增幅较大的集中在机器人和边缘操作系统、物联网操作系统、嵌入式操作系统和智能网联汽车操作系统等领域。

3.1.4 调研总览

在桌面操作系统方向上，依然是在Linux内核基础上的系统自主构建，以麒麟软件和统信软件为代表的第一梯队，在产品成熟度上取得长足进步，同时两家企业不约而同增加了自己对上游开源社区的贡献度，截止2024年3月，两家企业Linux内核贡献均超过100个，并先后于年内宣布建立“根社区”，暨不再依赖其他发行版做分支，而是进行操作系统的自主构建，这是Linux发行商技术能力提升的一个重要里程碑。同时在信创市场的引领下，麒麟和统信已经建成了百万级的国产操作系统应用软件生态，海量的应用软件为操作系统的国产化进程提供了巨大助力。2023年，国资委评选的十大国之重器，开放麒麟入选，也是中国开源软件首个国之重器。桌面操作系统未来发展是一个一直存在争议的话题，事实上尽管它是PC时代的产物，但同样面临着移动互联网时代和人工智能时代的冲击。麒麟和统信为代表的国产操作系统阵营，短期要解决替代的需求，同时与时俱进，提出下一代操作系统的新的定义和新的功能特征。据悉，两款国产桌面操作系统可以有效的与智能手机系统链接，并可以通过容器或类似ARC技术直接运行安卓APP，但我们认为下一代操作系统在工作场景、生活场景和娱乐场景都将深度的与人工智能技术集成，在人机交互、分布式计算、分布式存储、算力资源等领域均有巨大的发展潜力。微软公司作为最成功的桌面操作系统供应商，其当下的AI战略值得我们企业广泛借鉴。2023年9月统信率先发布UOS AI产品，现已覆盖90%主流开源大模型与AI框架。

在手机平板类操作系统方向上，得益于中国物联网产业的高度和手机平板等智能终端制造业的高度成熟，与桌面操作系统相比，这类操作系统普遍拥有更大量的实际用户、日活用户。除少量基于chromium和Linux+QT的技术架构外，基于AOSP的深度定制版依然是此类操作系统的主流架构。值得关注的一点是在过去一年中openHarmony正在成为此类操作系统的第二个选择，其在物联网类的上升势头更加明显，技术生态和影响力大幅提升。华为手机回归后，新一代的HamonyOS确实让业界眼前一亮，除了原有的软总线、混合内核等技术特征之外，其尝试在安卓生态之外，自建技术生态的战略发展路线着实让业内普遍为之振奋同时也为其担忧。一个技术生态的发展往往伴随着一个新时代的崛起，而技术生态又具有全球性和广泛性的特征，围绕单一市场是难以构建具备竞争力的新技

术生态。在生态构建者具备足够开源开放的基础上，同时还要有与之配套的长期发展战略和大量的资源投入，此外有与业界共享产业发展红利的胸怀，与生态伙伴共赢的商业模式，甚至还要抓住人工智能时代技术更迭的历史机遇。国内其他手机制造商对Harmony技术生态的态度可以作为一把标尺，也是其生态发展第一个要突破的瓶颈。

云和服务器类操作系统方向则呈现出openEuler、OpenAnolis和OpenCloudOS三驾马车持续领跑。作为华为云、阿里云和腾讯云的战略型核心基础设施，三款产品不论是在开放性、先进性和可持续发展性上均领先其他同类操作系统产品，但此类操作系统产品的同质化也是最高的（目前三家企业正在CAC的直接指导下加强技术合作，逐步在核心软件包和系统升级等方面开始统筹规划）。在所有收录的此类系统中，我们可以看到越来越多操作系统发行商正在将上游社区从Debian、RHEL/CentOS开始向openEuler和OpenAnolis转移，这一现象得益于近期国内兴起的“根社区”能力建设，这个方向无疑是正确的。自主可控操作系统，必须从自主构建开始，在开源操作系统生态中，如何提升自身的技术影响力，这其实是一个长期开源战略，自主构建只是第一步。

物联网操作系统方向则呈现出多种结构并存的百花齐放局面。一部分是基于Linux内核的衍生版（如OpenHarmony），也有一部分是基于RT-Thread内核的衍生版（如XiUOS矽璞等），此外近期基于OpenHarmony的衍生版在智能机器人和智能家居领域日渐增多。一直秉持自主研发的翼辉信息边缘操作系统EdgerOS近两年也取得了不俗的市场成绩。此类操作系统产品中还包括了许多智能网联汽车中的智慧座舱类系统。与前述几类系统最大的差别，是此类操作系统产品的硬件平台千差万别，因此系统定制化的要求较高，这也导致了开源操作系统的碎片化生态在此显而易见。对功耗、性能、工况、计算、存储和通信能力的差别短期内依然无法形成清晰产业格局，智能家居市场也成为目前看得见但却赚不到钱的“鸡肋”操作系统市场，尽管OpenHarmony正在成为越来越多此类系统的上游社区，但其技术生态影响力和市占率还远远不够，此类操作系统“小碎散”的现状短期难以扭转。

汽车类操作系统方向在智能网联汽车时代则呈现出三个应用场景驱动下的三类汽车操作系统，分别是车辆控制操作系统、智慧座舱操作系统和自动驾驶操作系统。另外由于智能域融合的发展趋势，涌现了对于嵌入式虚拟化操作系统的需求，在智能座舱、融合网关、舱驾合一的多个场景中都需要虚拟化来实现算力、外设资源的灵活调配。对于车辆控制操作系统当前依然是国外操作系统的天下，国内的东软睿驰、翼辉信息、睿赛德、中电科普华等少数企业有上车交付的案例，在智慧座舱和自动驾驶领域近年来涌现出了几十个技术产品，但同质化极其严重，可以说一车一系统，每个汽车主机厂都

在努力构建自己的品牌，大体上只有三个技术架构，第一个是“Linux深度定制”对标特斯拉的汽车操作系统（值得一提的是特斯拉在超级充电、自动驾驶等诸多领域的深度开放开源已经成为业界典范）；第二个架构是Linux+AOSP主要被国内的造车新势力和一级技术供应商所采纳，这类架构的特点是更侧重于操作系统本身的出发点来定义系统；第三个架构是QNX+AOSP，这也是目前国内最常见的技术架构，主要被国内传统汽车主机厂所采用，这类架构的特点是更侧重于汽车安全体系的出发点来定义系统，因此QNX成为了不二之选。汽车从早期机械时代、电控时代到如今的智能网联时代，汽车操作系统也正在从分布式、域控制向集中控制转化，因此此类操作系统尚未形成MacOS和Windows、iOS和Android一样的确定的两级产业结构，如何利用这个窗口期把握机遇抢占一个席位，值得我们深思。目前华为、斑马智行、中汽创智、小米、翼辉信息、中瓴智行、国汽智联、中科础石等均在摩拳擦掌，快速迭代产品。但此类系统最终仍需要通过“上车”这一关键环节来验证自身系统的安全性和可靠性，相较于对云和服务器的操作系统的国家级单位斡旋下的统一协作，汽车行业现阶段更需要行业主管部门和权威机构的介入，避免重复投资和产品同质化，避免“小碎散”局面在这个领域重现。其中，中瓴智行的RAITE Hypervisor嵌入式虚拟化产品精准定位于智能化域融合的基础设施，兼顾了资源隔离及高效复用、异构功能安全等级OS的生态兼容，已在多个主机厂实现了“上车”量产。

嵌入式类操作系统方向则根据其对操作系统的实时性要求不同，呈现出强实时和弱实时两个方向。强实时操作系统主要应用在国防军工、火箭卫星、大飞机、轨道交通、电力能源、工业制造、汽车电子、医疗设备、先进制造等对“实时性、稳定性、安全性”要求苛刻的领域，目前市场主要被VxWorks和其国内衍生版垄断，国内发展势头较好的有翼辉信息原创SylixOS，承担起我国关键产业的安全可控和持续发展。弱实时操作系统主要应用在智能穿戴设备、物联网设备、智能家居设备、工业互联网设备等行业，目前市场上Linux的定制发行版为主流方案，国内睿赛德RT-Thread、中移物联OneOS以及华为OpenHarmony发展势头相对较好。国内操作系统产品呈现出行业聚集的分布特点，即某个产品在某行业具有较高市场占有率，且很难在其他行业保持市场优势。此外中国电子、中国电科、中国航天、中国航空、中国船舶、中国兵器等央企所属的科研院所也分别拥有自己的嵌入式操作系统品牌，相对独立的服务于自身业务体系。

特种应用类操作系统方向则呈现出快速发展迹象，特别在WRZZ、TZZZ等场景下的操作系统产品快速迭代，在武器装备领域多个平台也涌现出更具竞争力的操作系统产品。

3.1.5 发展趋势

根据本次面向全社会的操作系统产品征集和调研，中国操作系统行业发展呈现出以下几个特点和趋势。

操作系统定义正在外延。随着云计算、物联网、人工智能和智能网联汽车等时代的来临，操作系统定义正在不断被扩大，虚拟化、容器、多模态感知、AI算法和业务中间件正在不断被操作系统发行商集成到操作系统内部，这一现象是新时代对操作系统不断诠释的正常演进，也造成了操作系统家族的泛化和产品细分。

操作系统产品的战略地位持续提升。自操作系统产品被定义为信息化核心基础设施之后，其在产业生态和供应链系统中的地位就在不断攀升。曾经普遍被质疑的操作系统商业模式随着时间的推移已经不再是业内人士口诛笔伐的焦点。现今，我国操作系统产品正在从产业政策驱动的时代向用户市场驱动的时代迈进。操作系统产品以其特有的技术生态核心地位正在成为各行业志存高远的科技企业构筑其核心竞争力的重要途径，成为企业长期竞争战略的重要组成部分。

操作系统产品同质化严重，生态碎片化加剧。操作系统产品因其基础性、广泛性和先进性的特点成为开源生态的重要组成部分。也正因如此，我国操作系统产品普遍通过继承或引用部分开源技术构建自身产品，这是同质化的根源之一。其次目前国内操作系统研发团队规模较小、成立时间短，只能将有限的研发资源优先投入到系统的定制化上，对上游开源社区和系统核心模块的投入不够，于是选择直接“继承”，导致了对系统结构框架、核心软件包以及部分强依赖软件包无法自主维护更新，软件包自研比例低，且相对解耦，如各种设备驱动程序，系统架构支持，运行时系统（这样做可以积累自身工作，避免主系统升级后，前期一些自研工作由于无法兼容和继承，而被破弃用）。同质化是生态发展的一个大忌，简单说就是低水平重复，浪费了大量的时间和资源。我国已经拥有超过二百多款操作系统产品，同质化的产品导致操作系统生态碎片化加剧，“小碎散”的局面依然广泛存在，因此产品的同质化是生态碎片化的一个重要原因。通过本次调研发现，产品同质化在所有操作系统类别中广泛存在，近年来在桌面类和服务器类中有逐步改善的迹象，但在物联网类和汽车类和嵌入式类中，产品同质化和生态碎片化的迹象还在加剧。

应用场景细分促使操作系统定制化成为趋势。碎片化的另一个原因是新时代对操作系统产品给出了更加丰富的应用场景。这里有两个重要趋势，一个是软硬一体化趋势，这里包括物理硬件也包括虚拟硬件，当我们选择一个在线服务、一款智能手表或者一台智能滴灌系统时，其供应商定制优化的系统将直接开始服务。可以试想 iPhone 手机刷一个鸿蒙系统？是的，这个想法已经过时了。另一个

就是智能化趋势，在物联网类和汽车类中，通过集成多模态交互功能和自动驾驶AI算法使得操作系统在该场景下拥有高度智能特征，甚至一些自主决策能力。这两个趋势或在未来一段时间内持续引领操作系统的发展方向。

操作系统产品研发的社区化和国际化持续推进。开源无国界，操作系统产品本身就是国际化合作的产物，是开源社区的研发成果，主要创新力来源于社区。我国操作系统产品的发展有赖于国际开源社区的几十年持续的开放、共享、交流和协作。但我们在开源文化底蕴培养开源社区治理运营上的进展则落后于我们下载开源代码的速度。2006年Linux-2.6.12内核中中国开发者仅贡献了50个软件包，如今中国开发者对内核社区的累计贡献已经在所有国家中排名第一（自Linux-2.6.12至Linux-6.8历时18年的98个版本累计数据），在华为、联发科、阿里、瑞芯微、中兴通讯、龙芯中科、VIVO、腾讯、字节跳动、成都蓉蓉、百度、中国移动、麒麟、酷派、OPPO、统信等一批企业的带领下，中国操作系统企业的社区化和国际化已经取得了一定的成绩。同期，这些企业创建的操作系统开源社区在国际上尚不具有影响力。希望未来中国企业能够建立可以比肩Kernel、Debian、Ubuntu、Fedora、CentOS这样具有国际影响力的操作系统社区，引领中国操作系统产业进入国际领先行列。目前，RT-Thread和Deepin是目前具有国际影响力较大的操作系统社区，openEuler和OpenHarmony以及OpenAnolis发展势头迅猛，希望未来可以获得全球开发者和国际市场的认可。

3.1.7 发展建议

针对本次调研所呈现的操作系统行业发展现状，我们提出以下三点建议：

- （1）我国操作系统产业的发展不仅需要政策引导、更需要获得市场引导和资本引导，才能进一步发挥我国的规模优势，调动企业研发的主动性和积极性，加强各操作系统团队间的深入合作，降低产品同质化、杜绝反复投资，将生态碎片化控制在可控范围内。
- （2）加强操作系统生态共建，鼓励各操作系统团队积极加强与上游开源社区的技术交流，提升自身影响力；加强与供应链相关技术社区的深入合作，补链强链，避免技术生态割裂带来的供应链安全隐患。
- （3）加强国内操作系统开源技术社区和用户群组的专业化运营和全球化建设，早日实现以开源社区驱动的操作系统的产品创新机制，把我们的“根”社区建成全球操作系统的“根”社区。

3.2 数据库产业发展报告

3.2.1 全球数据库发展态势

2023-2024年间, 在全球范围内, 数据库产业继续保持快速增长态势。云数据库服务的普及和开源技术的成熟推动了市场的扩大。根据国际数据公司(IDC)的报告, 全球数据库市场规模在2023年达到数百亿美元, 预计到2024年将继续保持两位数的增长率。

随着数字化转型的深入, 大数据、云计算、人工智能等技术的融合应用, 对数据库的需求激增。国际市场上, 云原生数据库服务成为主流趋势, 各大云服务提供商如AWS、Azure、Google Cloud等纷纷推出自家的云数据库服务, 强调弹性和成本效益, 推动了数据库即服务(DaaS)的快速发展。此外, 时序数据库、图数据库等专业领域数据库也因物联网(IoT)、社交网络分析等应用的普及而快速增长。

从2023年到2024年, 全球数据库发展态势体现在以下几个方面:

云数据库服务的增长: 随着云计算的普及, 云数据库服务继续扩大市场份额。云服务提供商如Amazon Web Services (AWS)、Microsoft Azure、Google Cloud Platform (GCP)等不断推出新的数据库服务和功能, 满足不同行业和应用场景的需求。

开源数据库的崛起: 开源数据库因其灵活性和成本效益受到市场的欢迎。PostgreSQL、MySQL、MariaDB、TDengine等开源数据库持续发展, 社区活跃, 贡献者众多, 功能不断增强。

分布式数据库的发展: 为了应对大数据和高并发的挑战, 分布式数据库技术得到快速发展。这些数据库通过分布式架构提供水平扩展能力、高可用性和容错性。

时序数据库和NoSQL数据库的创新: 特定类型的数据库如时序数据库(用于处理时间序列数据)和NoSQL数据库(如文档数据库、键值存储、宽列存储等)继续创新, 满足特定应用场景的需求。

人工智能和机器学习集成: 数据库开始集成人工智能和机器学习技术, 以提供智能查询优化、自动性能调优、预测性维护等功能。

- **数据库与大模型协同发展:** 大模型和数据库的相互融合在目前的大数据时代中正发挥着重要作用, 同时也为高效智能的数据管理与服务带来前所未有的机遇和挑战。
- **安全性和合规性:** 数据安全和隐私保护成为数据库发展的重要考虑因素。数据库提供商加强了数

据加密、访问控制、审计和合规性支持，以满足GDPR等法规要求。

- 多云和混合云策略：随着企业采用多云和混合云策略，跨云数据库服务和工具的发展，帮助企业更灵活地管理数据和应用。
- 自主可控数据库的发展：在全球政治经济形势变化的背景下，一些国家和地区推动自主可控的数据库技术发展，以减少对外部数据库技术的依赖。
- 数据库即服务（DBaaS）：数据库即服务模式为用户提供了更加便捷、灵活的数据库部署和管理方式，帮助用户降低运营成本，提高效率。
- 国际合作与竞争：数据库领域的国际合作与竞争加剧，跨国公司在全球范围内寻求合作机会，同时也面临来自本土创新企业的竞争。
- 市场规模的扩大：全球数据库市场规模持续增长，根据市场研究报告，预计在2023至2024年间将保持稳定的增长率。

综上所述，全球数据库市场在这一时期内呈现出多样化、技术创新、安全性加强和多云支持等发展特点。随着技术的不断进步和市场需求的增长，数据库产业预计将继续快速发展。

3.2.2 中国数据库发展态势

中国数据库市场同样呈现出高速增长态势。政府政策的大力支持，如“新基建”、“数据要素X”等计划的推进，为数据库产业发展创造了有利条件。据《2023中国大数据产业发展指数报告》显示，中国大数据产业持续加速，头部城市的大数据技术标准推进和应用示范项目显著增多，带动了数据库产业的蓬勃发展。

市场规模方面，预计到2024年底，中国数据库市场总规模将达到数百亿元人民币，年复合增长率超过20%。随着企业对数据管理和分析需求的增加，数据库作为数据基础设施的核心，其市场需求将持续增长。

2023年至2024年，中国数据库发展态势主要表现在以下几个方面：

快速增长的市场规模：随着数字化转型的加速推进，中国数据库市场规模持续扩大。企业对数据管理和分析的需求不断增长，推动了数据库软件和相关服务市场的快速发展。

- 国产数据库的崛起：国产数据库厂商加大研发投入，推动技术创新，市场份额逐步扩大。在政府政

策的支持下，国产数据库在金融、电信、政务等领域的应用不断深入。

- 云数据库服务的普及：云服务提供商如阿里云、腾讯云、华为云等，加速布局云数据库服务，提供包括关系型数据库、NoSQL数据库、时序数据库等多种云数据库产品，满足不同行业和应用场景的需求。
- 开源数据库的发展：开源数据库在中国市场上得到广泛应用，社区活跃，贡献者众多。国产数据库厂商积极参与开源项目，通过开源协作推动技术创新。
- 技术创新与融合：中国数据库厂商持续推动技术创新，包括云原生数据库、分布式数据库、时序数据库等。同时，数据库技术与人工智能、大数据、物联网等新技术的融合也在不断加深。
- 安全性与合规性重视：数据安全和隐私保护成为数据库市场的重要考虑因素。国产数据库厂商加强数据加密、访问控制、审计和合规性支持，以满足国家法律法规的要求。
- 行业应用的深化：数据库产品在金融、电信、政务、制造、医疗、教育等多个行业得到广泛应用，推动了这些行业的数字化转型和智能化升级。
- 市场竞争格局变化：市场竞争激烈，国产数据库厂商通过技术创新和市场拓展，与国际数据库厂商形成竞争。同时，不同数据库厂商之间也在寻求合作，共同开拓市场。
- 投资与并购活动：随着市场的快速发展，投资和并购活动增多。资本的注入为数据库厂商提供了更多的发展资源，有助于加速技术创新和市场扩张。
- 政策支持与引导：中国政府出台多项政策支持国产软件和信息技术服务的发展，鼓励企业采用国产数据库产品，推动数据库产业的自主创新和国产化进程。

综上所述，中国数据库市场在2023年至2024年间呈现出快速增长、技术创新、国产化推进、行业应用深化、市场竞争加剧等发展态势。国产数据库厂商加大研发投入，推动技术创新，市场份额逐步扩大。特别是在金融、电信、政务等领域，国产数据库的应用不断加深。随着数字化转型的不断深入，预计中国数据库市场将继续保持快速发展的势头。

3.2.3 开源数据库发展情况

(1) 全球开源数据库发展趋势

2023年至2024年，全球开源数据库的发展情况呈现以下几大趋势和亮点：

技术创新与融合趋势显著

- **云原生与Serverless**: 在云原生方面的技术创新主要体现在容器化和微服务架构的广泛应用, 这些技术使得数据库能够更灵活地部署在云环境中, 提高其可扩展性和可移植性。同时包括对Serverless计算的支持, 通过自动扩展和按需付费的模式, 能够更经济高效地处理波动的工作负载, 降低运维成本。此外, 技术创新还表现在与云服务提供商的深度集成, 例如提供专用的数据库服务和工具, 简化了数据库的配置、监控和维护, 同时增强了与云生态系统的互操作性。
- **HTAP技术**: 混合事务与分析处理 (HTAP) 技术在开源数据库中的应用增加, 使得单一数据库系统能同时处理事务操作和分析查询, 减少了数据迁移的复杂性和延迟, 提高了系统在处理复杂混合负载时的性能和效率。
- **多模数据库**: 多模方面的主要技术创新是增强了对多种数据模型的支持, 包括关系型、键值对、文档、图形和时间序列数据, 以满足不同应用场景的需求; 不断提升了多模数据库的查询能力, 通过统一查询语言或API简化了对不同数据模型的访问, 使用户能够在单一数据库系统中高效地处理和分析混合数据类型。
- **与AI融合发展**: 开源数据库通过机器学习算法优化数据库性能, 例如自动调整查询优化器的参数以提高查询效率; 集成自然语言处理 (NLP) 能力, 使用户能够以自然语言查询数据库, 从而简化数据访问并提高用户体验。此外, 通过支持向量类型和向量计算能力, 可以高效地存储、索引和搜索向量数据, 与大模型 (如大规模推荐系统或自然语言处理模型) 结合使用时, 能够加速机器学习任务和提高相似性搜索的性能。

开源生态日益繁荣

- **开源项目库建设**: 全球开源软件产品库, 特别是“全球数据库产品库”的建设得到推进, 成为连接全球开发者、促进技术共享与合作的重要平台。
- **社区活跃度上升**: 开源数据库社区的活跃度持续提升, 开发者和用户之间的互动更加频繁, 促进了技术创新和问题解决的效率。
- **标准化与合规性**: 随着开源数据库在关键业务领域的广泛应用, 对标准化和合规性的重视程度提高, 以确保数据安全和隐私保护。

市场接受度与应用广度

- 行业应用深化：开源数据库在金融、电信、电商、医疗、物联网等行业中的应用案例不断增多，尤其是时序数据库如IoTDB、TDengine、InfluxDB等，在物联网和工业互联网领域展现了强大潜力。
- 企业采纳度提升：越来越多的企业，包括大型企业和初创公司，选择开源数据库作为其核心系统的一部分，以降低TCO（总拥有成本），提升技术灵活性。

中国开源数据库的崛起

- 本土项目发展：中国本土开源数据库项目如TDengine、openGauss、TiDB、OceanBase、StoneDB、IvorySQL等，不仅在国内市场获得广泛应用，也开始在国际舞台上竞争，展示出中国在开源数据库领域的创新能力。
- 政策与市场双重推动：中国政府对开源软件的鼓励政策，加上市场需求的强劲增长，为中国开源数据库的发展营造了良好的外部环境。

2023年至2024年，全球开源数据库的发展在技术创新、生态建设、行业应用等方面均有显著进展，显示出开源技术在数据库领域的活力与潜力。随着技术的不断成熟和应用场景的拓宽，开源数据库将继续在数字化转型中扮演重要角色，推动数据管理技术的革新与发展。

(2) 中国开源数据库发展现状

2023年到2024年，中国开源数据库的发展现状显示出以下几个关键特点：

- 市场规模迅速增长：中国数据库市场规模持续扩大，预计到2024年，市场规模将超过400亿元人民币，反映出企业对于数据库，尤其是开源数据库解决方案的强烈需求。
- 技术创新与突破：在数据库核心技术方面，中国厂商和开源社区持续取得创新成果，如在分布式处理、云原生架构、HTAP（混合事务分析处理）、多模数据库等方面均有显著进步。头部项目如TDengine、openGauss、TiDB、OceanBase、OpenTenBase、IvorySQL、IoTDB、PolarDB等，在性能优化、安全性、易用性、兼容性上不断迭代升级。
- 开源生态日益成熟：开源数据库的社区建设成效显著，以openGauss和TDengine为例，openGauss社区生态在国内外均得到了快速发展，吸引了大量开发者、企业和研究机构参与，共同推

动技术创新和生态繁荣；TDengine开源5年时间在GitHub发展了23.1kStar，建立了数十个用户群，全球用户实例数达到了59万，得到了众多开发者和企业的支持。

- 行业应用深化：开源数据库在金融、电信、政府、电商、医疗等多个行业实现了规模化应用，如openGauss在国计民生行业规模商用，TiDB在电商、金融行业的广泛应用，loTDB在我国成飞、宝武等头部企业深度应用，TDengine在电力能源等传统行业中的广泛应用，显示出国产数据库在关键领域的替代能力和竞争力。
- 政策支持与引导：政府对国产化、自主可控的重视，以及对开源技术的积极态度，为开源数据库的发展提供了良好的政策环境和资金支持，推动了开源数据库在国家重大工程和项目中的应用。

（3）中国开源数据库的应用领域

- 金融行业：金融领域对数据处理的高要求促进了openGauss、PolarDB、TiDB、OceanBase等数据库在金融行业的广泛应用，特别是在交易系统、风控分析、客户关系管理等核心业务场景中。
- 智能制造与物联网：随着工业互联网的发展，开源数据库如loTDB、TDengine等在设备数据收集、实时分析、预测维护等方面发挥重要作用，支持智能制造和智慧城市等领域的数据处理需求。
- 电商与新零售：高并发、大数据量的电商平台倾向于采用分布式数据库如PolarDB、TDSQL等，以应对促销高峰期的流量洪峰，同时支持个性化推荐、库存管理等复杂业务逻辑。
- 医疗健康：在医疗健康领域，数据库如SequoiaDB被用于电子病历管理、医疗影像存储、基因数据分析等，支持PB级数据处理，保障数据安全与隐私。

总体而言，2023至2024年的中国开源数据库发展聚焦于技术创新与行业深度融合，旨在满足不同行业对数据处理速度、灵活性、安全性的多样化需求，推动数字经济的高质量发展。

（4）中国开源数据库典型代表

2023年到2024年间，中国最具代表性的开源数据库在技术特点与研发创新方面展现出了显著进步和特色，以下是一些突出的方面：

TiDB：TiDB作为分布式NewSQL数据库，主打水平扩展、强一致性和HTAP能力。它能够在单个数据库中同时进行联机事务处理(OLTP)和联机分析处理(OLAP)，适合需要实时数据分析的应用场景。在这两年中，TiDB团队致力于优化分布式事务处理性能，增强SQL兼容性，以及提升自动化运维工具的

智能化水平。此外，TiDB也在持续探索更高效的索引策略和查询优化技术，以适应更复杂的数据查询需求。

OpenTenBase: OpenTenBase开源分布式HTAP数据库—100%兼容MySQL和PostgreSQL生态，由腾讯于2023年捐赠给开放原子开源基金会孵化及运营的开源数据库项目，具备高扩展性、商业数据库语法兼容、分布式HTAP引擎、多级容灾和多维度资源隔离等能力，目前已经成功应用于金融、证券、医疗、航天、交通等诸多行业的核心业务系统。OpenTenBase与腾讯云数据库TDSQL商业版内核云上云下同源，目前在运行实例公有云有30多万个实例，是拥有最大“练兵场”的国产开源数据库。

TDengine: TDengine 核心是一款高性能、集群开源、云原生的时序数据库，专为物联网、工业互联网、电力、IT 运维等场景设计并优化，具有极强的弹性伸缩能力。同时它还带有内建的缓存、流式计算、数据订阅等系统功能，能大幅减少系统设计的复杂度，降低研发和运营成本，是一个高性能、分布式的物联网、工业大数据平台。当前 TDengine 主要提供两大版本，分别是支持私有化部署的TDengine Enterprise 以及全托管的物联网、工业互联网云服务平台 TDengine Cloud，两者在开源时序数据库 TDengine OSS 的功能基础上有更加强，用户可根据自身业务体量和需求进行版本选择。

PolarDB: PolarDB是由阿里云推出的一种云原生数据库，它支持存储与计算分离的架构，可以独立扩展计算或存储资源。PolarDB高度兼容MySQL、PostgreSQL等，提供了高性能和低成本的数据服务。2023至2024年间，PolarDB在自动化运维、智能诊断、弹性扩缩容等方面取得了显著进展，尤其是在数据库的自我优化和故障预测上，运用了更多的机器学习技术，提升了用户体验和系统稳定性。

openGauss: openGauss聚焦于关系型数据库内核的优化，致力于提供高性能、高可靠性的数据库服务，并且在SQL执行引擎、并发控制、存储管理等方面有独到之处。openGauss在内核层面的创新包括但不限于多核CPU的高效利用、分布式事务处理的优化、以及对容器化环境的更好支持，同时加强了对最新硬件特性的适配，以提升整体系统性能。

IvorySQL: IvorySQL不仅完美继承了PostgreSQL的先进特性，并在此基础上增强了对Oracle数据库的兼容性，为用户从传统数据库迁移提供了平滑的路径。在性能与安全性上，IvorySQL进行了深度优化，确保在高并发环境下也能提供稳定的高速数据处理能力，同时内置了严格的访问控制与数据

加密机制，充分保障企业数据资产的安全。针对云时代需求，它无缝集成云原生技术，支持容器化部署与自动化运维，成为构建现代化云基础设施的理想选择。

loTDB: loTDB是一款清华大学研制的开源工业物联网时序数据库管理系统，采用端边云协同的轻量化架构，支持一体化的物联网时序数据收集、存储、管理与分析，具有多协议兼容、超高压缩比、高通量读写、工业级稳定、极简运维等特点。它打破了关注计算中心后台的传统数据库管理系统边界，形成了端上采集、边缘汇聚、云中处理的上行链路，以及云中训练、边缘推理、端上执行的下行链路，该软件在BenchAnt国际权威榜单中六项指标位列世界第一，成为工业物联网高效运行的基础软件。

StoneDB: 产品对标Oracle HeatWave，使用MySQL的用户无需进行数据迁移，通过StoneDB可以实现TP+AP混合负载，交易和分析性能最高可达MySQL的2倍和400倍，通过AP增强到自主可控的TP，瞄准大量MySQL信创升级+替代市场。

这些开源数据库在技术特点上各有侧重，研发创新则集中于提升性能、扩展性、安全性和智能化运维，反映了中国在数据库技术领域紧跟国际前沿，同时结合本土市场需求进行创新的动态。通过不断的研发和优化，这些数据库正逐步缩小与国际领先数据库的差距，甚至在某些领域展现出超越态势。

3.2.4 数据库技术发展

(1) 数据库技术未来趋势

数据库技术的未来发展趋势可以从以下几个关键方面进行概括：

- **云原生与Serverless数据库:** 云原生数据库将继续发展，以更好地适应云环境，包括自动扩展、高可用性和与云平台的深度集成。Serverless数据库将成为一种趋势，允许用户无需管理底层基础设施，根据实际使用量按需付费，极大简化了数据库运维并降低了成本。
- **分布式与弹性架构:** 随着数据量的爆炸性增长，分布式数据库技术将更加成熟，提供横向扩展能力，以处理大规模数据集并保证高可用性和高性能。存储计算分离、Shared Nothing架构等设计将促进资源的弹性扩缩容和高效利用。
- **HTAP (混合事务分析处理):** HTAP数据库将事务处理和分析处理能力融合，减少数据移动，提高

数据处理的实时性和效率，适应快速决策支持和实时分析的需求。

- **AI与机器学习集成**：数据库系统将更紧密地集成人工智能和机器学习技术，一方面用于自动化运维、性能优化、查询预测和智能数据管理，另一方面直接利用SQL语句进行模型训练与推理部署的能力，提升数据库系统的智能水平和自适应能力。
- **数据安全与隐私保护**：随着数据保护法规的日益严格，数据库将内置更为先进的加密技术、隐私保护机制和合规性工具，确保数据在传输、存储和处理过程中的安全。
- **多模态数据支持**：数据库将更好地支持半结构化、非结构化数据以及新兴的数据类型，如图像、视频等多媒体数据，以及图数据和时序数据，满足多样化的数据处理需求。
- **开源数据库的崛起**：开源数据库因其灵活性、成本效益和强大的社区支持而持续增长，中国本土的开源项目如openGauss、TiDB、IoTDB等将进一步提升技术水平和生态建设，增强国际竞争力。
- **硬件技术的融合**：数据库技术将与新型硬件技术（如持久内存、GPU加速、量子计算等）更紧密结合，以实现更高性能和更低延迟。
- **边缘计算与物联网应用**：随着物联网设备的普及，边缘数据库技术将得到发展，以支持低延迟数据处理和局部自治，满足物联网应用的即时分析需求。
- **可持续性和绿色计算**：数据库技术也将考虑能源效率和环境影响，发展节能优化策略，减少碳足迹，适应绿色计算趋势。

这些趋势综合体现了数据库技术正朝着更灵活、智能、高效和安全的方向发展，以适应数字化转型和数据密集型应用的不断增长需求。

未来，数据库技术将更加侧重于云原生化、智能化、自动化管理，以及跨平台兼容性。随着5G、边缘计算的普及，边缘数据库、混合云数据库解决方案将成为新的热点。同时，AI与数据库的深度融合，如自动调优、智能预测等，将进一步提升数据库系统的性能与效率。

（2）中国数据库产品技术研发与创新

2023年至2024年，中国数据库头部厂商如阿里云、腾讯云、华为云等在国内数据库市场占据领先地位，纷纷加大研发投入，推动技术创新。阿里云的PolarDB、腾讯云的TDSQL、华为的GaussDB等产品，均在分布式处理、实时分析、自动化运维、数据安全等方面取得显著进步。其中，PolarDB的云原

生架构设计，以及TDSQL在金融级高可用和安全性上的突破，代表了中国数据库技术的前沿水平。中国数据库产品在技术研发与创新方面展现出了显著的进展和活力，具体表现在以下几个方面：

- 云原生与分布式技术的深度融合：中国数据库厂商和研究机构深入探索云原生架构，开发出更加轻量级、弹性的数据库产品，支持容器化部署、微服务架构，以及Kubernetes原生集成，提高了数据库在云环境下的部署效率和运维管理能力。分布式数据库技术持续优化，不仅关注水平扩展能力，还深入到数据一致性、容错恢复、跨数据中心复制等高级特性，满足大规模数据处理和高可用性要求。
- HTAP技术的成熟应用：HTAP（Hybrid Transactional and Analytical Processing）技术在中国数据库产品中得到广泛应用，实现了事务处理（OLTP）和分析处理（OLAP）的融合，减少数据孤岛，提高数据处理效率，尤其在金融、电商、物流等领域展现出巨大价值。
- AI与数据库融合创新：通过引入机器学习和人工智能技术，中国数据库产品在查询优化、资源调度、故障预测等方面实现了智能化，提高了性能和运维效率。例如，使用AI模型预测数据库负载，自动调整资源配置，以及利用智能算法优化SQL查询执行计划。
- 安全与隐私保护技术强化：针对数据安全和个人隐私保护的严格要求，中国数据库产品加强了数据加密、动态数据脱敏、细粒度访问控制等安全特性，确保数据在传输、存储、处理各环节的安全合规。
- 数据库内核优化与创新：在数据库内核层面，中国数据库厂商进行了大量优化工作，如改进存储引擎、优化事务处理机制、提升并发处理能力等，以提高数据库的性能和可扩展性。部分数据库产品还实现了对新型硬件（如GPU、SSD）的充分利用，进一步加速数据处理速度。
- 国产数据库生态构建：国产数据库如openGauss、TiDB、PolarDB、TDSQL、HGDB、IoTDB等加大了生态建设力度，通过举办开发者大会、技术培训、社区贡献奖励等方式，吸引国内外开发者参与，推动了技术交流与合作，形成了较为完善的生态体系。
- 行业定制化解决方案：针对特定行业需求，中国数据库厂商推出了一系列行业定制化解决方案，如金融行业的高并发交易处理、医疗健康领域的数据隐私保护、智能制造的数据实时分析等，这些解决方案提升了数据库在垂直领域的应用价值和市场竞争力。
- 标准化与国际化：中国数据库产品和技术积极参与国际标准化进程，推动数据库接口、协议、安全

标准的统一，增强了中国数据库在国际市场上的互操作性和认可度。

综上所述，2023年至2024年中国数据库产品在技术研发与创新方面取得了显著成就，不仅提升了技术实力，也加强了国际影响力，为中国乃至全球数字经济的发展提供了坚实基础。

3.2.5 中国数据库产品需求与市场竞争

(1) 国内企业对国产数据库发展的需求

国内企业对国产数据库发展的需求主要体现在以下几个方面：

- **自主可控与信息安全：**随着数据安全和国家信息安全意识的提升，企业对国产数据库的需求首先源自对数据自主可控的强烈需求。企业希望采用国产数据库能够降低依赖外国技术的风险，确保数据不被非法获取或利用，满足国家安全法规和行业监管要求。
- **行业适配与定制化服务：**不同行业有着特定的业务需求和数据处理特点，如金融行业对交易处理速度和数据一致性要求极高，政府机构则更重视数据的加密与隐私保护。因此，企业期待国产数据库能提供高度行业定制化解决方案，满足特定业务场景下的需求。
- **性能与稳定性：**在数字化转型加速的背景下，企业对数据库的性能和稳定性有着极高的期望。这包括支持高并发访问、快速数据处理、高可用性设计等，以确保业务连续性和用户体验。
- **云原生与混合云支持：**随着云计算的普及，企业对能够无缝集成云环境的数据库需求日益增长，尤其是云原生数据库和能够支持混合云部署的解决方案，以便企业灵活选择云服务提供商，实现资源的最优配置。
- **易用性与运维简便性：**为了降低IT运维成本，企业倾向于选择易安装、易维护、界面友好的数据库产品。此外，对自动化运维工具、智能化监控和故障预测的需求也在增加，以减轻运维负担，提高效率。
- **生态兼容与互操作性：**企业希望国产数据库能够与现有的IT系统和应用程序良好兼容，包括对主流编程语言的支持、与主流中间件和服务的集成能力，以及与其他数据库的迁移工具和接口兼容性，确保平滑过渡和升级。
- **成本效益：**在保证性能和安全的前提下，成本控制也是企业考量的重要因素。国产数据库应提供具有竞争力的价格优势，以及灵活的付费模式，如按需付费、订阅制等，帮助企业有效控制IT支出。

综上所述，国内企业对国产数据库的需求综合了技术先进性、行业适应性、安全合规性、易用运维、成本效益等多方面因素，期待国产数据库能够提供全面、可靠、高效的数据管理解决方案，支持企业的数字化转型和长期发展。

（2）中国数据库产品发展过程中的主要问题及对策

2023年至2024年，中国数据库产品发展面临的主要问题集中在核心技术竞争力、生态体系成熟度、国际化水平、专业人才短缺，以及数据安全与合规挑战上。为应对这些挑战，中国数据库产业正采取一系列积极对策：加大科研投入，聚焦核心技术突破，以提升产品性能与创新能力；构建开放生态，促进国际合作，丰富周边工具和服务，增强用户体验；实施国际化战略，提升国际认证与服务水平，拓宽海外市场；通过校企合作、专业培训等手段，加强人才培养和引进；并强化数据保护技术，确保产品符合国内外安全合规要求。通过这些综合措施，中国数据库产业正努力克服现有障碍，加速技术进步和市场拓展，以满足国内企业对自主可控、高性能、安全合规数据库产品的需求，促进数字经济的健康发展。

（3）中国数据库产品竞争态势与市场占有情况

2023年至2024年，中国数据库产品的竞争态势呈现出以下几个特点：

- 国产替代加速：受到政策支持和信创（信息技术应用创新）政策的驱动，国产数据库产品在政府、金融、电信、能源等关键行业的应用明显加速，替代进口数据库产品的趋势持续增强。企业更加重视数据安全与自主可控，国产数据库厂商因此获得了更多的市场机会。
- 技术创新与产品迭代：为应对激烈的市场竞争，中国数据库厂商不断加强技术创新，推出支持云原生、分布式架构、AI融合、高可用性、高性能的新产品或版本，以满足企业对数据处理速度、灵活性、扩展性的更高要求。
- 生态体系构建：构建完善的生态系统成为国产数据库厂商提升竞争力的关键。厂商纷纷通过与上下游合作伙伴协作，开发工具链、中间件、应用软件的适配，以及提供专业服务来丰富和完善自身的生态系统。
- 市场竞争加剧：随着市场参与者增多，既有传统数据库厂商的转型，也有新兴数据库公司的崛起，以及互联网巨头的布局，市场竞争日趋激烈。在Gartner魔力象限等国际权威评估中，中国数据库厂商的表现逐渐亮眼，显示出国产数据库在全球市场中的地位提升。

- 细分市场分化：数据库市场进一步细分，不同类型的数据库（如关系型数据库、非关系型数据库、时序数据库、图数据库等）针对特定应用场景优化，满足不同行业和业务需求，形成差异化竞争格局。

- 国际化尝试：部分领先的国产数据库厂商开始探索国际化道路，通过参加国际展会、寻求海外合作伙伴、通过国际安全认证等方式，提升国际知名度和市场份额。

- 服务与解决方案能力提升：除了产品本身，国产数据库厂商还通过提升售后服务、提供定制化解决方案、优化用户培训和支持体系，增强客户黏性，提升市场竞争力。

总体来说，中国数据库市场竞争格局正处在快速变化之中，国产数据库厂商在政策、技术、市场等多重因素的推动下，正逐步打破国际品牌长期占据主导的局面，形成多元化、高质量的竞争态势。

对于2023年至2024年国外与国产数据库在中国市场的占有率，直接的具体数字没有详细信息，但可以根据现有趋势和行业分析进行一般性概述：

- 国产数据库市场占有率：国产数据库在这段时间内市场占有率持续上升。2022年，国产数据库软件市场规模已达到204.32亿元，占数据库市场86.42%的份额。这里需要注意的是，这个数字可能特指数据库软件产品规模而非整个数据库市场的占有率，且包含了各种类型的数据库产品。尽管如此，这反映了国产数据库在特定领域的强劲增长。随着政策推动、技术进步和国产替代需求的增加，预计2023年至2024年，国产数据库的市场占有率会进一步增长，特别是在政府、金融、电信等关键行业。

- 国外数据库市场占有率：相比之下，虽然国外数据库厂商如Oracle、IBM、Microsoft等在中国市场依然保持一定的份额，尤其是在高端市场和传统企业领域，但面对国产数据库的崛起，其市场份额面临压力。政策导向和国产化趋势促使部分市场转向国产数据库，因此国外数据库厂商的市场占有率可能呈现缓慢下降趋势。不过，这些国外品牌依旧依靠其技术优势、品牌影响力和全球服务体系，在某些领域维持着稳固的地位。

综上所述，2023年至2024年期间，国产数据库市场占有率持续增长，而国外数据库市场占有率则可能因国产替代趋势而略有下降，但具体数字需要参照最新的行业报告或市场调研结果。

（4）中国数据库发展重要事件

2023年至2024年间，中国数据库企业及厂商发生了一系列重要的事件，这些事件彰显了中国数据库

行业的发展活力和技术进步，主要包括：

标准参与与荣誉认可：

贝格迈思（深圳）技术有限公司在2023年荣获CCSA TC601委员会“2023年度优秀成员单位”，深度参与数据库相关国家标准与行业标准的制定，凸显了国产数据库厂商在行业标准制定中的积极作用。

产品创新与技术突破：

- 华为云数据库RDS在2023年底至2024年初期间，发布了新版本，强调了成本节省与查询性能提升，展示了在数据库产品创新方面的持续努力。
- 腾讯云TDSQL在2023年中国PostgreSQL数据库生态大会上荣获“最佳数据库产品”奖，体现了其在分布式数据库解决方案和行业应用实践中的领先地位。
- 2023年IoTDB在国际数据库测试榜单中存储占用、写入吞吐率、读取吞吐率、查询延迟、操作成本等关键指标排名世界第一。IoTDB的开放文件格式TsFile在2023年底被Apache社区接受为顶级项目。

国产数据库厂商名录更新与成就展示：

2023年，多家国产数据库厂商如人大金仓、瀚高软件、南大通用、达梦数据库、神舟通用、PingCAP等，继续在数据库自主研发、技术创新、市场拓展等方面取得显著成绩，展示了国产数据库企业蓬勃发展的态势。

行业应用与案例推广：

根据第一新声调研，目前党政领域的数据库国产替代率高达80%，已经基本处于替换的尾声阶段，八大行业由于对国产数据库的稳定性、迁移难度、运维难度等方面要求更高，所以整体替换率不高，金融行业非核心系统处于40%左右，能源行业不足15%，医疗、制造、教育等多个行业甚至不足5%。

- 在党政领域，党政机关已基本完成国产数据库的应用系统改造和达梦启云数据库云服务系统成功中标雄安新区综合数据平台建设工作。典型应用案例包括：达梦启云数据库云服务系统成功中标雄安新区综合数据平台建设工作项目、GaussDB支撑陕西财政搭建分布式数据库平台。

- 金融行业中，银行业的非核心系统，国产数据库替换比例接近50%，核心系统替换比例15%左右；证券和保险业，非核心系统国产数据库使用比例均低于30%，核心系统低于20%。典型应用案例包括OceanBase支持工商银行对公(法人)理财系统的分布式改造、海量数据库支持中华联合人寿核心团险国产化改造项目。
- 电信业方面，各运营商从集团到省公司提出明确要求，到2022年要求去O比例达到50%，到2027年完成全替换。典型案例包括：南大通用山东移动经分系统改造项目、科蓝软件中国联CBSS集中计费系统应用。
- 能源行业目前处于早期阶段，但超80%的能源企业TOP10中采用国产数据库。典型案例包：括瀚高数据库中标国家电投集团自主可控数据库采购项目、崖山数据库中标深圳燃气数据库国产化项目。
- 医疗行业自2023年底，医疗行业各系统如电子病历、疾病瀚高数据库支撑山东大学附属儿童医院离线应急系统:检测、手术麻醉、数字认证、办公系统等开始进行国产数据库替换升级。典型案例包括：瀚高数据库支撑山东大学附属儿童医院离线应急系统、达梦数据库上线厦门大学附属成功医院核心系统。
- 制造业国产数据库替换整体滞后于党政和金融等领域，部分企业仍处于非核心系统“试点替换”，核心系统处于观望阶段。典型案例包括：海量数据库支撑比亚迪打造全国智造核心系统、瀚高数据库支撑江西中烟工业有限责任公司数据中台底座建心系统。

3.3 人工智能产业研究报告

3.3.1 人工智能产业概览

(1) 人工智能技术及应用领域概述

人工智能(AI)是一门融合计算机科学、生理学和哲学的综合性学科,旨在通过模拟和拓展人类的智能来解决复杂问题。人工智能技术的发展与计算机科学行业紧密相关，并在诸多领域展现广阔的应用前景。

随着科技不断推进，人工智能(AI)技术逐步渗透到各行各业,成为推动智能化发展的重要力量。人工智能技术旨在模拟人类的智能思维和行为模式，通过大数据、机器学习等技术手段,赋予机器自动执

行复杂任务和决策的能力,实现高效智能的人机交互。下面将从传统人工智能和人工智能大模型两个方向展开介绍。

传统人工智能技术与应用发展进展

目前,传统人工智能技术主要包括六大核心应用领域:自然语言处理、机器视觉、机器学习、专家系统、智能机器人和推荐系统。

自然语言处理(NLP) 是人工智能的关键技术之一,研究如何使机器能够理解、处理和生成自然语言,实现流畅的人机语言交互。NLP技术通过分析文本语义、语法和语用特征,广泛应用于语音识别、机器翻译、情感分析、智能问答等领域。

NLP在人工智能领域占据重要地位,是研究的重点方向,旨在探索人机通过自然语言实现高效互动的技术手段。目前,NLP研究主要集中在语言学、数据处理和语言工程等方面。加拿大蒙特利尔大学的TAUM-METE机器翻译系统是该领域较为成熟的应用。NLP技术的应用范围十分广泛,包括机器同声传译、智能客服、聊天机器人等。

自2000年以来,语音识别技术日趋成熟,识别准确率可达95%以上,如百度的语音识别系统。NLP技术在客户服务领域应用普遍,移动通信时的智能语音问答就是一个典型案例。随着人工智能的持续发展,人机对话功能将更加普及,这有助于解决人口老龄化等社会问题。

机器视觉(Computer Vision) 是人工智能领域的又一重要技术,旨在让机器模拟并具备人类视觉系统的功能,实现对图像和视频的理解与分析能力。机器视觉技术可广泛应用于图像识别、目标检测、人脸识别、行为分析等领域。

机器视觉技术致力于通过计算机模拟人类视觉系统,代替人眼识别物体、确定物体位置和运动状态。在技术实现层面,机器视觉通常包括三个关键步骤:目标检测(图像预处理、图像分割)、目标识别(特征提取、目标分类、匹配判断)和行为识别(模型构建、行为分析)。

机器视觉技术在不同细分领域的发展存在一定差距。目前,生物特征识别是该技术最为成熟的应用场景,在考勤、安防等领域得到了广泛运用,如指纹识别、人脸识别、虹膜识别等。然而,在物体场景识别方面,由于物种类别繁多、外观特征多样,技术尚不十分成熟。

近三年来,机器视觉行业取得了突飞猛进的发展,行业应用也在不断扩展。除了传统的3C电子、汽车和半导体产业外,机器视觉技术还逐渐渗透到了医疗健康、物流仓储、农业等多个领域。

未来,随着算法的不断优化和硬件设备的持续升级,机器视觉技术必将在更多场景展现出强大的性能,为人机交互、智能分析等领域带来革新性体验。

机器学习(Machine Learning) 作为人工智能的核心技术之一,主要研究如何利用数据和经验不断优化和提升机器的性能表现和决策准确性。机器学习技术通过构建并训练模型算法,使机器能够自动从海量数据中学习,获取经验知识,并基于学习成果对未知情况做出预测和决策。

机器学习技术已被广泛应用于广告推荐、金融风控、医学诊断、智能交通等诸多领域,展现出巨大的商业价值和社会意义。它有望为人类社会带来前所未有的创新体验和生产力提升。

机器学习可谓是人工智能领域的核心所在,解决了智能系统"学习"的关键问题。机器学习通过让机器利用现有信息持续学习、更新知识库,并产出智能化的输出结果。

我们不难发现,机器学习技术已经突破了人类智能的部分极限。以“AlphaGo”大战李世石的例子,虽然在第一局中人类下Chess大师李世石击败了AlphaGo,但在第二天的对局中,AlphaGo凭借自身在夜间与自身进行的上百万局对弈练习,达到了连李世石这样的人类顶尖大师都无法企及的水平。机器学习让机器在"学习"的过程中突破自我,实现了人类难以做到的知识积累和运算能力。

第四种人工智能技术是专家系统(Expert System),它基于模拟人类专家的决策过程和知识表达方式,通过编码专家知识和推理规则,实现对复杂问题的智能化解决。专家系统广泛应用于医疗诊断、法律咨询、工业控制等领域,帮助人类在专业领域内做出准确决策。

第五种是智能机器人(Intelligent Robot) 技术。智能机器人是指具备感知、认知和控制能力的机器人系统。它可借助传感器获取环境信息,对信息进行感知和理解,并根据既定需求和目标做出决策并采取行动。智能机器人已广泛应用于制造业、服务业、医疗护理等领域,为人类提供更加便捷高效的服务体验。

智能机器人与普通机器人的区别在于,智能机器人具有相对发达的“大脑”,能实现自主控制。一般而言,智能机器人需具备三大要素:感知外部环境的感觉能力、与环境交互的运动能力,以及对信息进行处理并做出反馈的思考能力。

目前,智能机器人的应用领域主要集中在工业机器人、服务机器人和产业链。工业机器人应用于汽车焊接、电子装配、食品搬运等。服务机器人则渗透至医疗手术、家居服务、娱乐教育等领域。而在产业链层面,机器人本体、控制系统、传感器等都是发展的重点。

在医疗健康领域，智能机器人的应用较为成熟。2016年在日本就有一例，医生诊断无计可施时，智能机器人“沃森”在十几分钟内阅读了两百万页相关医疗文献，并给出了救治方案，最终挽救了那名患者的生命。除此之外，智能机器人也正在远程医疗、康复护理、智能假肢、微创手术等领域发力，为医疗健康事业贡献力量。

在家庭服务领域，智能机器人的应用迎来黄金增长期。当下应用较多如扫地机器人、智能音箱、家庭安全监控设备等。以家居清洁为例，现有的智能家务机器人已能实现自主游走、垃圾清理、地面清洁等功能。

最后的人工智能技术是推荐系统。推荐系统通过分析用户的历史行为和兴趣爱好，预测用户可能感兴趣的内容或商品，并为其提供个性化推荐服务。该技术已广泛应用于电商、社交媒体、音视频等平台，提升了用户体验和商业价值。

总的来说，传统人工智能在自然语言处理、计算机视觉、机器学习、专家系统、智能机器人和推荐系统等六大领域展现出广阔的应用前景。随着相关技术的不断创新，人工智能必将为人类社会带来更多便利与可能。

人工智能大模型技术与应用发展进展

大型语言模型（以下简称LLM）被广泛认为是迈向人工通用智能（AGI）的关键路径之一。2022年1月30日，一个名为ChatGPT（即聊天生成预训练变换器）的LLM问世，短短两个月的时间内，该模型就吸引了超过一亿的活跃用户。LLM凭借其强大的适应能力和逻辑推理能力，能够执行聊天对话、撰写邮件、创作诗歌、编写代码以及制定商业提案等多种任务，其问世标志着AGI走向新的里程碑。图1便生动展示了自2013年以来LLM在发展历程中所取得的进步，以及模型参数规模的壮观扩张。

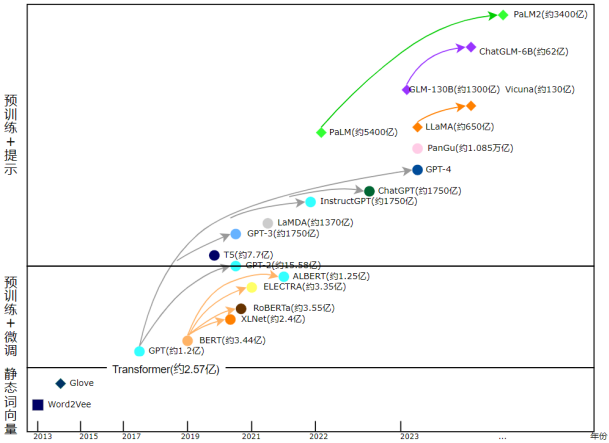


图1 大语言模型的发展历程及参数规模

在Word2Vec和GloVe这类词向量表示模型问世前，对文本语义的理解主要停留在浅层次。这些模型实现了对文本语义进行分布式向量化表示，但却受限于它们的静态特性，无法有效应对一词多义的语境挑战。为了克服这一局限并提升对自然语言的深层理解能力，引入了包含自注意力机制的Transformer架构，该架构通过编解码器堆叠方式，使语义和任务特征的提取能力远远超越了传统深度学习模型如卷积神经网络。

目前市场上的主流语言模型基本上都是建立在Transformer的基础之上，这些模型虽然基础架构同质，但构建方法和应用路径则是多样化的。例如，BERT利用了Transformer的编码器来实现双向的特征捕获；GPT则通过Transformer的解码器来实现单向的文本生成；而T5模型则融合编码器与解码器的优势，实现从文本到文本的转换任务。这些模型的提出和发展不仅极大地推进了自然语言处理技术的边界，也为AGI领域带来了新的思路和机遇。

随着BERT和GPT系列模型的问世，大型语言模型的架构逐渐演化为两种主流技术路径，即预训练加上微调（Pre-training + Fine-tuning）和预训练加上提示（Pre-training + Prompting）。BERT模型代表了预训练加微调的方法，它首先通过大量无标注数据，诸如维基百科条目和网络新闻，进行了充分的自我监督预训练。随后通过特定任务的有标注数据进行微调，以在特定场景下提升模型的表现。而GPT系列代表了预训练加提示方法，利用大规模无标注数据先行自监督预训练，而后通过零样本（One-shot）或少样本提示（Few-shot Prompting）来帮助模型更好地理解特定任务。这种提示通常通过人工或自动生成的自然语言提示来实现，使得在参数规模足够大时，模型具备少数样本甚至零样本情况下的文本生成能力。

在大型语言模型发展初期，BERT模型的微调模式受到更多关注，这主要是由于BERT的双向特性让其在理解类任务上的表现超越了GPT的单向模型。因此涌现出了BERT派生模型，如XLNet、RoBERTa、ALBERT以及ELECTRA。随后，OpenAI研究团队推出了参数更大、数据集更庞大的GPT-2和GPT-3。正如表1所显示，GPT-2的数据量和参数规模是GPT的大约十倍，达到了15亿个参数；而GPT-3的参数规模进一步扩大到了1750亿。当大型语言模型参数日益增长时，它们在进行任务微调时所需的成本也随之增大。因此，GPT-3及其后继模型在应用到下游任务时，通常不进行梯度更新或进一步微调，而是倾向于使用少数样本提示，提供有限的上下文实例进行推理，亦即所谓的上下文学习（In-Context Learning）。这得益于更高的数据质量和更大的数据规模，使得大型语言模型在少数样本学习方面展现出惊人的涌现能力。

GPT-3，作为拥有数千亿个参数的模型，展现出了其雄厚的知识处理能力。然而，它所依赖的概率性

文本生成机制也可能产生非真实、无用甚至有潜在害处的输出。这一挑战并非通过简单扩大模型规模所能解决。于是，InstructGPT和ChatGPT应运而生，它们的目的是使大型语言模型的输出更加符合人类意图。InstructGPT以GPT-3为基础，而ChatGPT则基于GPT-3.5架构。值得一提的是，GPT-3.5在训练时结合了GPT-3和Codex模型的数据，并以InstructGPT为基础，经过人工标记数据的有监督微调而成。2023年，GPT-4被提出，标志着大型多模态模型的新阶段，能够处理文本和图像的多模态信号。

作为自回归语言模型和提示学习方法的代表，GPT系列得到了众多企业和教育机构的广泛支持。例如，谷歌在2022年发布了LaMDA和PaLM。LaMDA在GPT-3的架构基础上进行了改进，通过人工标注数据的微调和模型的外部知识利用，显著增强了模型在安全性和事实性方面的能力。继LaMDA之后，谷歌在2023年3月推出了Bard模型，这是经过LaMDA微调的一款模型。PaLM是Google Pathways系统训练下的首个超大规模语言模型，参数扩展至5400亿，并且大幅提升了多语言性能。在2023年5月，PaLM 2作为与GPT-4对等的模型被推出，尽管其训练参数大幅减少，但它展现出了比PaLM更优越的多语言和推理能力，并且在计算效率上更高，在一些任务上实现了与GPT-4相当乃至更优异的性能表现。

2023年2月，Meta公司发布了开源的大型语言模型LLaMA (Large Language Model Meta AI)。LLaMA模型的特点是只利用公开可获得的数据进行训练，并做了更多的标注以得到相对小型的模型，这让它在运行效率和资源消耗上优于其他模型。斯坦福大学在LLaMA-13B模型的基础上研发了Vicuna模型。凭借约300美元的训练成本，Vicuna-13B模型已达到了ChatGPT和Bard模型90%以上的效能。而英伟达/微软的Megatron-Turing模型，华为的PanGu- α 和PanGu- Σ 等十亿级参数规模的大模型也都采纳了GPT-3的架构，并在此基础上作了改进。

除了BERT和GPT系列模型架构，清华大学的研究团队提出了一种名为GLM (General Language Model) 的模型架构，它利用自回归的填空方式进行学习。基于GLM，清华大学在2022年10月推出了一个拥有1300亿参数的中英双语预训练大型语言模型GLM-130B。在此基础上，2023年3月推出了ChatGLM-6B模型，这是一个支持中英双语的对话语言模型，特别优化了中文问答和对话的生成能力。通过对大约1万亿中英标识符进行训练，并配合监督微调、反馈自助和人类反馈强化学习等技术，ChatGLM-6B模型已能产生极为符合人类喜好的结果。此外，还有一些基于ChatGLM的特定领域应用模型也相继被研发出来，例如DoctorGLM等。

(2) 人工智能大模型的局限及发展趋势

尽管如ChatGPT这样的大型语言模型标志着聊天机器人的进步，从简单的任务自动化走向更深层的智能，但在技术和安全方面还面临着诸多挑战和风险。

技术局限:

- 输出存在真实性和偏见问题。虽然通过RLHF算法优化了模型的输出以更贴近人类意图，大型语言模型仍有可能生成违背常识、带有偏见和歧视的不当回复。多样化的评价指标也难以确保模型在各方面达到均衡，仍有真实性存疑的风险。
- 缺少实时自学能力。大型语言模型作为一个封闭系统，缺乏可解释性，基础模型也难于根据新数据更新知识库。尽管“预训练+提示”模式能够使得模型输出更接近人类意图，但这种基于提示的学习方式不能及时地吸纳新知识，也不适合处理长篇复杂或高度专业的语言结构内容。尽管可以通过微调模式来集成新知识，这种方式效率较高、成本较低，但可能会造成所谓的“灾难性遗忘”。
- 依赖大量高质量数据。大型语言模型的表现取决于庞大的、高质量的训练数据。尽管RLHF算法降低了部分数据标注成本，但仍依赖于大量的高质量标注数据。训练数据中的语言风格、用语习惯及标注者的个人偏好都会对模型产生重大影响。另外，对于资源较低的语言，优质的训练数据较少，使得模型在这些语言上的表现难以达到英语处理的水平。
- 语言生成的趋同性和单一化。虽然大型语言模型的文本生成速度快、准确率高，但它们在语言多样性和富有表现力方面尚有不足。与谷歌翻译等翻译工具相比，在处理常见语言时表现出色，但对于资源较少的语言则表现较差。如ChatGPT的英文训练数据量远超其他语言，导致其多语言处理能力有限。
- 需要大量计算资源。大型语言模型的部署和运行依赖于高性能的计算资源，这意味着巨大的能源消耗和环境成本。因此，模型的轻量化和性价比更高的计算平台的开发迫在眉睫。

安全风险:

- 数据隐私问题。大型语言模型可能在未经充分授权的情况下处理用户数据，存在个人隐私泄露或滥用的风险。
- 信息安全问题。模型可能在缺乏可靠来源支撑的情况下生成看似可信的内容，对判断能力不足的

用户构成潜在误导。同时，由于用户群体庞大，这些模型具备一定的社会舆论引导和动员能力，可能引起信息传播的安全隐患。

- 数据伦理问题。模型输出可能反映训练数据中存在的偏见，包含歧视性或违反社会准则的内容。此外，面对不同文化和社会背景，模型可能无法做出恰当的反应，增加了跨文化伦理的风险。
- 降低违法犯罪成本。技术的操纵性使得它可用于帮助犯罪行为，如通过仿真社交环境诱导透露敏感信息，从而便利了犯罪活动。
- 鉴于大型语言模型在技术层面的局限和安全层面的风险，其发展不仅需要技术创新，也需要政策监管和安全上的保障。

技术改进方向:

- 从“大规模”到“轻量化”模型的转变。研究揭示了模型的参数量、数据量、以及训练过程的计算量与效果之间的缩放关系。虽然品质随规模增加而提升，但这也带来了对计算资源的巨大需求。为了降低成本，当前趋势是采用模型压缩技术，如量化、剪枝和稀疏化等以减小模型规模和计算需求。
- 从“单一模态”到“多模态”模型的发展。大型语言模型正逐渐引入更多的多模态信息，例如图像和声音，以更好地与人类认知对齐，满足多样化的应用需求。这需要更丰富和高质量的训练数据，有助于提升模型性能。如GPT-4和KOSMOS-1，已经开始将多模态特性整合进模型中。
- 从“通用模型”到“领域特定垂直模型”。通用大语言模型无法充分满足个性化需求且训练成本高昂，未来的发展将转向针对特定行业以深入解决行业痛点的垂直领域模型。这类模型的训练成本、资源消耗均小于通用模型，并且能够提供更加专业、深入的行业解决方案。例如，Bloomberg News基于GPT-3为金融领域定制了模型，谷歌推出了医学和安全两个领域的专用模型Med-PaLM 2和Sec-PaLM 2。

政策调整与安全监管:

人工智能监管成为全球政府关注的重点，当前缺乏针对生成型人工智能的专门监管机制和完善法规。

- 构建人工智能风险评估框架。明确大型语言模型如ChatGPT的潜在风险指标并通过评估来发现及处置隐患，加强监管和制定风险防范计划。

- 强化多维监管和责任落实。涉及技术、行业和主体信用的多层面监管措施，例如算法备案制度、基于风险分级的行业规则和法律責任机制，以及企业和个人对应的安全义务和法规。
- 推动法律适用和行业自律。将现有的法律如《网络安全法》、《数据安全法》、《个人信息保护法》扩展适用到大型语言模型，同时建立行业法律与规范，促进自律机制的形成。
- 大型语言模型如ChatGPT在人工智能的发展中起到了标杆作用，不仅展示了出色的通用性能和逻辑推理能力，也带来了转变AI研究的范式和应用模式。未来的发展趋势将是模型轻量化，多模态能力的扩展，以及向垂直领域模型的演变。与此同时，政策监管将必不可少，确保技术的创新发展与社会安全、伦理和法律框架同步前进。

(3) 人工智能产业链及主要参与者

人工智能产业链是硬件软融合、多主体参与的复杂生态系统，主要包括基础层、技术层、应用层三层，如图2所示：

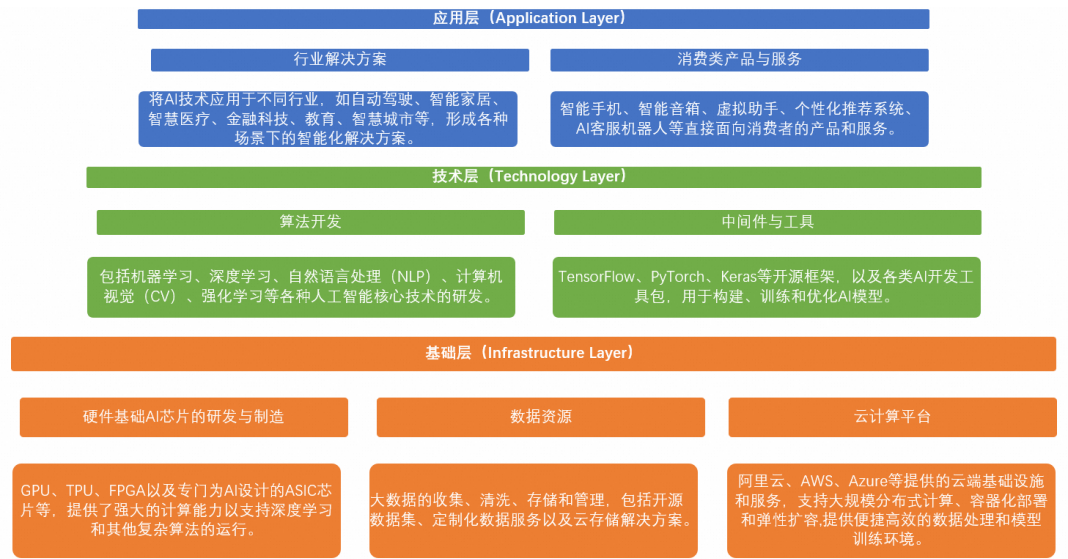


图2 人工智能产业链三层结构

基础层 (Infrastructure Layer) 主要包括算力和数据。包括硬件基础、数据资源和云计算平台三个组成部分。硬件基础主要为AI芯片的研发与制造企业，为AIGC (Artificial Intelligence Generated Content, 即人工智能生成的内容) 提供算力支撑深度学习和算法运行。数据资源同样属于基础层，

包括数据的收集、清晰、存储和管理，以语言类大模型为例，“token”是最基本的数据处理单元，需要将数据处理为token，经过编码后变成一个唯一的数字ID，这样模型就可以通过数学运算来处理原本难以解析的自然语言文本。模型通过学习海量token及其在上下文中的关系，建立起对语言的理解和生成能力。云计算同样也是重要的基础设施，通过云操作系统，支持大规模分布式计算、容器化部署和弹性扩容，为企业和开发者提供便捷高效的数据处理和模型训练环境。

技术层（Technology Layer）主要包括算法和AI训练工具。机器学习、深度学习、自然语言处理（NLP）、计算机视觉（CV）、强化学习等各种人工智能核心算法。中间件与工具：如TensorFlow、PyTorch、Keras等开源框架，以及各类AI开发工具包，用于构建、训练和优化AI模型。

应用层（Application Layer）主要包括解决方案和智能终端。将AI技术应用于不同行业的解决方案通常是一系列应用和功能的综合组织，如智慧城市解决方案。另一个重要部分为面向终端消费者提供的服务，包括智能手机、智能音箱等硬件，也包括虚拟助手、AI客服机器人等直接面向消费者的产品和服务。

此外，产业服务在人工智能产业链中也占据着重要地位，包括三类内容：一是咨询服务，即提供AI战略规划、项目实施等咨询服务的企业。随着人工智能技术与产业应用的融合程度提升，这类服务需求激增。二是教育培训机构，主要为AI人才提供职业教育培训和在线课程平台，支持从业者获取最新职业培训。三是标准制定与测试认证方。参与制定AI相关国际国内标准的组织，以及提供AI产品性能测试与认证的服务机构。整个产业链还包括政策法规制定者、科研机构、投资机构等多元角色，共同推动人工智能技术的发展和落地。随着模型技术发展，对不同模型性能的测评已经成为驱动模型技术发展的重要力量。

国际上对大模型（尤其是大型预训练语言模型）的测评主要聚焦于其在多个任务上的表现、泛化能力、资源效率、可解释性等多个维度。以下是一些通用且重要的评测标准和内容：

跨任务评估：

GLUE Benchmark (General Language Understanding Evaluation): 针对自然语言理解任务设计的一套基准测试，包括文本蕴含、相似性判断、问答等多种任务类型。

Big-Bench: 谷歌推出的大规模语言模型评估基准，包含众多涉及常识、逻辑推理、数学能力和世界知识等复杂任务。

Multi-Task Evaluation Suites: XTREME, 跨越多种语言和任务类型的评估基准, 测试模型在翻译、句法分析、命名实体识别等多个任务上的性能。

MMPT (Multilingual and Multitask Prompted Training): 考察模型在多语言环境下对多种任务的零样本和少样本学习能力。

Foundation Model-Specific Benchmarks: C-EVAL, 这是一个针对基础模型 (Foundation Models) 的多层级多学科评估套件, 旨在全面评估模型在广泛NLP任务上的高级能力。

此外, IEEE大模型标准工作组也在制定相关标准, 包括《大规模深度学习模型评估框架及流程》, 涵盖了从模型性能到部署实践等全方位的评估内容。随着技术的发展, 新的评估基准和指标不断出现, 旨在更全面地反映大模型的真实效能和潜在风险

从参与者类型来看, 人工智能产业主要参与者包括硬软件及算法研发、产业应用、监管部门及协会组织等组成, 也有覆盖链条较长、环节较多的头部互联网平台:

- **科研机构:** 如大学、研究所等, 负责前沿理论研究和关键技术突破。
- **科技巨头:** 如谷歌、微软、亚马逊、阿里巴巴、百度、腾讯等, 拥有强大研发实力, 覆盖从基础技术到应用产品的全产业链条。
- **创业公司与独角兽:** 专注于某一细分领域的AI技术或应用, 例如商汤科技、旷视科技、优必选等, 在人脸识别、图像识别、语音识别、具身智能等领域具有影响力。
- **芯片制造商:** 如NVIDIA、英特尔、AMD、华为海思等, 提供专门针对AI计算需求的硬件支持。
- **数据服务商:** 提供数据标注、数据清洗等数据预处理服务的企业。
- **垂直行业企业:** 将AI技术集成进自身产品和服务的传统行业企业, 如医疗影像诊断公司、自动驾驶汽车制造商等。

此外, 政府、行业协会、投资者和用户也是产业链中的重要组成部分, 共同推动着人工智能产业的发展与创新。

(4) 人工智能产业市场规模及增长趋势

人工智能产业近年来在我国高速发展。据中国信通院测算, 2022年中国人工智能核心产业规模达

5080亿元，同比增长18%；初步统计，2023年人工智能核心产业规模达到5784亿元，增速达13.9%。2023年随着ChatGPT带来的生成式人工智能大模型发展，人工智能市场格局发生突变，生成式人工智能发展带来算力需求涌现。2023年11月3日，人工智能计算大会发布的《中国人工智能算力发展评估报告》显示，2023年我国人工智能算力市场规模将达到664亿元（IDC预测24年中国AIGC市场规模在14亿美元左右），同比增长82.5%，人工智能算力需求快速增长。从地区分布看，北京、杭州、深圳、上海、苏州排名前五，智算中心的建设是拉动地区实现人工智能发展的重要驱动力。未来中国人工智能市场发展要取决于应用普及速度。

当前从行业渗透率看，互联网行业作为人工智能技术应用的主阵地居于首位，电信、金融和制造业排名靠前。从应用场景看，知识管理、对话式应用、代码生成等是企业应用人工智能的主要场景。目前，67%的中国企业已经开始探索人工智能的应用机会或已经开始进行相关资金投入，中国的人工智能技术应用已经迈入加速阶段。预测到2027年之前，中国智能算力规模年复合增长率将达33.9%，同期通用算力规模年复合增长率为16.6%。IDC预测中国到2027年，人工智能市场总体规模将接近400亿美元。

在应用层，IDC调研显示，在未来一年半左右，所有企业都会把生成式AI当做一个企业发展的核心，预计接近40%的核心业务将会被生成式AI所影响，企业在AI方面的投资增长接近25%。

（5）开源在人工智能产业中的作用和现状

开源技术在计算机科学和技术的发展中扮演着至关重要的角色，从早期的Unix生态系统到如今的大规模AI和云计算，开源运动一直驱动着创新、协作和共享知识的进步。开源技术在每个阶段都显示出了其降低准入门槛、加快创新速度、保障软件质量和安全、以及促进社区合作与共享成果的独特优势，从而对计算机技术的发展产生了不可估量的影响。

对于人工智能产业发展，开源技术使得人工智能领域的研发不再是封闭的个体行为，而是全球开发者共同参与的过程。诸如TensorFlow、PyTorch、Keras等开源框架促进了深度学习和其他AI技术的迅速迭代和标准化，大大加快了整个行业的研发速度。企业和个人开发者可以免费使用开源AI工具和平台，降低了入门门槛和开发成本，尤其对于初创公司和科研机构来说，不必从零开始构建基础设施，可以直接在开源基础上进行创新。大量人工智能开源社区的繁荣鼓励全球范围内的开发者、研究人员和企业进行合作交流，通过贡献代码、文档、模型和案例，形成了一个高度活跃的知识分享网络，提高了整体技术水平和解决方案质量。开源环境为人工智能人才提供了实践和学习的平台，通过

参与开源项目，开发者能够提升技能，熟悉最新技术动态，解决实际问题，从而快速成长为符合市场需求的专业人才。开源促进了产业链上下游的整合与联动，如硬件厂商优化对开源框架的支持，企业围绕开源项目打造产品和服务，形成健康的生态系统。

引发本轮人工智能大模型浪潮的Transformer架构自2017年由Google的研究团队在论文《Attention is All You Need》中首次提出以来，彻底革新了神经机器翻译和其他序列转换任务的方法，取代了传统的循环神经网络（RNN）和卷积神经网络（CNN）在许多自然语言处理任务中的地位。Transformer通过自注意力机制解决了长距离依赖问题，并因其并行计算友好性而大大提高了模型训练效率，公开的理论研究推动了工程化实践。TensorFlow & PyTorch TensorFlow由Google Brain团队开发，PyTorch由Facebook团队开发，这两个深度学习框架的开源极大推动了AI研究和应用的进程。研究人员和工程师可以轻松访问先进的模型和算法，从而快速迭代新的AI解决方案。

2023年Meta推出的开源LLaMA模型（Large Language Model by Meta AI, 或简称Llama）提供了多种参数规模的语言模型供研究者和开发者使用，降低准入门槛，使得更多人能够参与到大模型的研发与应用中来，促进了整个领域的开放性和透明度。LLaMA模型在Transformer架构基础上可能进行了创新性的改进和优化，包括但不限于对注意力机制的改进、模型结构的精简或增强等，这些技术进步不仅提升了自身模型性能，也为后续大模型的设计提供了参考和借鉴。LLaMA系列模型开源了多个参数量级，从小规模到大规模均有布局，且在诸如对话场景等方面表现优异，这表明它们能适应不同的应用场景需求，有助于推动大模型在实际业务中的广泛应用。开源的LLaMA模型让较小规模的研究团队和公司也能获得高质量的语言模型资源，减少了对闭源商业大模型的依赖。

美国模型开源社区已经实现商业转化。Hugging Face人工智能开源社区专注于自然语言处理（NLP）领域，重点围绕NLP模型开源、模型库、工具和API开源，成立于2016年。最为人所知的是其开源项目Transformers库，这个库包含了一系列基于Transformer架构的预训练模型，比如BERT、GPT-2、RoBERTa、XLM-R等，极大地推动了NLP研究和应用的民主化进程，使得开发者无需从头开始训练模型就能快速实现诸如文本生成、问答、文本分类等多种任务。Hugging Face还运营了一个名为“模型_hub”的在线平台，用户可以在上面发现、分享和使用各类AI模型和数据集。该平台吸引了包括微软、Meta、OpenAI、英伟达等在内的诸多企业和研究组织参与，共同促进AI技术的开源和协同创新。随着社区的发展和影响力的扩大，Hugging Face公司也获得了资本市场的高度认可，截至2023年数据显示其估值达到了20亿美元，反映出开源模式在AI产业发展中的重要性和前景。

在中国，随着开放原子开源基金会等组织的成立和推进，国内开源生态环境得到了大力扶持与发

展,越来越多的企业和个人参与到开源AI项目中,提升了本土创新能力。截止至2023年底及2024年初,中国在大模型领域的开源活动相当活跃。截止到2023年底,阿里通义千问、智谱chatGLM、幻方量化deepseek-67B模型、零一万物的Yi-34B模型、百川的baichuan2-13B模型等陆续开源。阿里与中国计算机学会共建AI开源社区魔搭是目前国内最大的模型开源社区,共开源模型超4000个,阿里贡献了超千个行业最优模型,智源研究院、智谱、百川、澜舟等海淀研究机构和大模型企业积极参与社区贡献;汇集全球开发者超450万,累计提供免费算力超3000万卡时。百度星河社区包含飞桨PaddlePaddle深度学习框架等核心项目,旨在搭建起产学研用一体化的开源生态,推动国内AI技术的研发与应用。北京智源人工智能研究院作为国内领先的人工智能研究机构,同样积极推动开源工作。虽然没有具体的案例细节,但推测其可能通过开源研究成果、工具和平台等方式,来促进学术界与工业界的深度融合,培养高质量的人工智能人才,并帮助中小企业更快地采用和落地先进的人工智能技术。

(6) 人工智能开源面临的问题、风险和挑战

人工智能开源的确促进了技术进步和行业生态的繁荣,但也面临着多维度的问题、风险和挑战。

一是安全漏洞与滥用风险。一项2018年的研究表明,通过对GitHub上的数十万个Python项目进行扫描,发现许多流行的机器学习库存在可利用的安全弱点(来源:OWASP ML-Security Project)。开源AI模型容易被复制并用于恶意用途,比如生成伪造的内容,或者在未经适当安全措施下应用于关键基础设施,从而增加安全风险。

二是知识产权与商业利益冲突。尽管许多大型科技公司在推进AI开源项目,但他们在保护自己的核心技术知识产权方面仍然非常谨慎。例如,Google虽然在TensorFlow等项目上采取了开源策略,但在某些高级功能和最新研究成果上仍会有所保留,以维持自身在市场中的竞争优势。

三是模型偏见与伦理问题。2020年,MIT Media Lab的一项研究指出,语言模型如GPT-3在处理性别、种族和社会敏感话题时表现出显著的偏见(来源:《On the Dangers of Stochastic Parrots》)。

四是合规与监管挑战。随着GDPR等严格数据保护法规的出台,开源AI项目在处理个人数据时需要格外注意合规性。例如,欧盟在2021年强化了AI相关的法规草案,其中特别提到AI系统的可追溯性和透明度要求,这对开源AI提出了更高的合规门槛。

五是标准化与互操作性。在开源AI领域,不同的框架、模型和工具之间的互操作性一直是挑战。例

如，尽管PyTorch和TensorFlow都是开源的深度学习框架，但是它们之间的模型转换和代码移植并不总是无缝对接。为此，业界一直在努力推进统一的标准和中间件，如ONNX（Open Neural Network Exchange），以促进跨平台和框架的兼容性。

六是可持续发展与社区维护。开源AI项目在初期可能得到大量关注和支持，但后续的维护和升级却可能因资源有限而受阻。例如，很多开源项目由于缺乏持续的资金支持和技术投入，随着时间推移可能会失去活力，导致安全性更新滞后，或是不能及时适应技术发展趋势。

3.3.2 中国人工智能技术实力

（1）中国人工智能大模型发展历程

自从20世纪50年代人工智能技术诞生以来，已经历了从专家系统到机器学习，再到深度学习以及当前的大模型时期四个重要发展阶段。尽管前三个阶段因受限于数据、计算能力、应用场景和投资效果等因素而没有完全实现AI产业的商业化，大模型时代凭借其通用性、涌现性及强大的计算力，在商业应用成熟度上显示出巨大潜力。

- **专家系统阶段：**在早期，专家系统是人工智能领域的先锋，主要采用规则和推理方法来模仿专家知识。专家系统通过编码专家知识与规则，结合推理引擎来模拟解决问题的决策过程，但受限于构建知识成本高、表达困难和泛化技术能力弱，仅在商业化循环的数据及知识层面取得进展。
 - **机器学习阶段：**随后，基于数据的机器学习方法推动了人工智能的进步。其核心是从数据中识别出模式和规则，为计算机系统赋予智能行为。但彼时受限于适用场景较少、对训练数据的严重依赖及计算资源的限制，主要进步在“算法+算力”层次。
 - **深度学习阶段：**深度学习作为机器学习的子集，通过使用深度神经网络学习模式识别，利用多层的神经网络结构掌握高层次的特征表达，但因泛化能力和投资回报率不足，并未形成完整的商业闭环。
 - **大模型时代：**目前，我们迎来了大模型时代。大模型代表了预测模型的重要演进，它们拥有巨大的参数量、复杂结构，能够展现前所未有的通用性和创新能力。随着数据和算力条件的成熟及对多任务有更高的适应性，大模型提高了技术的商业潜力，并有望构成一个自足的商业闭环。
- 从技术可实现性和未来的广阔空间来看，越来越多的参与者投身于发展大模型技术，为AI商业化开拓更广阔的天地。大模型阶段最终可能实现人工智能技术的商业利用，成为AI产业化的一个重要转

折点。

自2022年12月ChatGPT-3.5问世以来，全球对超大规模语言处理模型的关注急剧上升。随后，中国高等院校等科研机构、科技巨头如百度、阿里巴巴、360和商汤科技等公司推出各式各样的大模型。至2023年6月，中国市场上涌现出大量的通用和领域特定大模型。目前，中国公开的大模型数量已超过300个，市场呈现出竞争激烈的多极格局，形成了一场被业内称为“百模大战”的竞争热潮。

在这个不断演变的大模型市场中，各家机构纷纷推陈出新，竞争激烈。2023年2月，复旦大学首发MOSS模型，北京大学随后推出了ChatExcel的测试版。3月，百度发布了文心一言，360推出了1.0版本智脑大模型，商汤科技推出了SenseNova，阿里巴巴推出了通义千问，昆仑万维发布了天工模型。4月，复旦大学开源了MOSS，同时，出门问问也发布了序列猴子模型。

接着不久，政策也开始响应市场热潮，北京和深圳相继出台了推动AI发展措施。6月，随着更多产品如百川智能的Baichuan-7B和科大讯飞的星火计划的加入，大模型行业的发展态势愈发显著。其中，10亿级参数规模以上的大模型数量已达79个，41个算法模型完成了备案。

随着时间推进到2023年下半年，新的参与者持续涌入，像华为、京东和网易等也加入了这场大模型的竞赛，带来更多更新的版本和应用，让市场竞争的格局日趋明朗。各大机构的新模型和更新迭代，如百川智能的Baichuan-53B、腾讯的混元大模型、蚂蚁金融的大模型以及上海人工智能实验室的书生系列模型，都是这一时期的杰出代表。

进入2024年，中国大模型领域持续蓬勃发展。各类大模型纷纷涌现，完成备案的模型数量不断攀升，显示出行业的高速增长和法规的积极跟进。各家企业不断推出新的升级版大模型，从华为的盘古升级版到商汤科技的SenseChat V4，再到抖音发布的AI构建平台Coze“扣子”，持续引领着行业的风潮。

总之，伴随着“百模大战”的激烈竞争，通用和垂类大模型的发展格局逐渐明朗，竞争日益激烈。这些模型在赋能行业转型升级的同时，中国的AI大模型产业正奔跑在走向更加智能的未来，迈向人工通用智能（AGI）时代的征程上。

（2）中国在人工智能芯片、系统等硬件算力层面的实力

随着ChatGPT的广泛流行，AIGC行业得到了行业界和公众的集中关注。大规模语言模型的拓展正推动着AI算力市场步入新纪元，积极构建起强健的AIGC运算基础架构。在复杂的系统工作中，大模

型的训练不仅是一项系统性挑战，而且也需要系统化的计算支撑。基于这一背景，我们开展了AIGC领域计算力的深入研究，力图在AIGC产业不断演进的过程中，把握并理解其计算力发展的各种动态。

在AIGC元年，计算力已成为驱动整个产业进步的核心动力，决定着技术的发展节奏和质量。AI的计算能力不仅局限于模型训练，它还支持多领域、多场景下的AI应用，催生出多样化的产品和服务新模式。

2022年年底，得益于ChatGPT的成功，我们见证了大量高参数规模的通用大模型相继问世。这些大模型的训练过程不仅需要处理千亿乃至万亿数量级的参数，还要经历上千GB的高质量数据，训练过程中的迭代次数进一步加剧了对高性能计算资源的需求。

AIGC产业的飞速发展，在很大程度上依赖于我们称之为“暴力美学”的强大计算资源，而具体的应用实践需要计算力在现实场景中的有效落地。AI技术实际应用包含两个关键环节：训练与推理。训练涉及通过海量数据构建AI模型，旨在满足特定需求，代表着AI技术的研发过程，进一步高级化的参数量对计算力需求极大。而推理环节则利用训练完善的模型，通过处理实际输入数据获得相应的结论，通常体现在日常应用中，关键是对实际应用场景数据吞吐量的处理能力。

AIGC领域的运算能力是一个工程化的产物，涵盖了从微芯片设计到综合资源服务的多个层面。随着未来大规模模型的产业应用不断深入，其所要求的系统工程变得更为复杂，中心任务是建立起既高效又稳定的计算力支撑平台。在这个过程中，成熟的算法技术、健全的数据链、配套的工具集以及广泛的生态系统构成了实现优化解决方案的关键要素。

在构建AI计算力平台方面，软硬件的兼容性以及调优性能的专业技巧是保证AI系统适应性和稳定性的重要保障。这些细节处理所需要的技术知识不是由单一参数或简单组件就能决定的，而是需要通过综合考虑多方面因素并应用深层次的工程专业知识的。

AIGC行业的快速扩展加速了AI芯片市场的增长，并促进了产品线的多样化。随着生成型人工智能技术的进步及其在各种场景中的应用，如数据中心、边缘计算基础设施以及终端设备等，对AI芯片的生产和应用需求也随之增加。展望未来，据预测到2027年AI芯片市场的规模将超过2023年的两倍，达到接近1194亿美元的高峰。这一显著增长反映了AI芯片在技术进步和市场拓展两方面的重要作用。

随着大模型的参数数量经历了指数级增长，训练这些模型的过程迫切需要依赖于由GPU组成的集群化计算资源。

在AI芯片的定义中，更为狭隘的理解是指那些为人工智能算法特别设计以实现加速的芯片。当前，随着深度学习算法在人工智能界的认可度和实际应用的快速增长，AI芯片主要被定义为那些专门为处理和训练大规模数据集以及执行推断任务而优化的芯片。

在AIGC时代，随着模型复杂度的增加，内存成为了计算力有效利用的关键决定因素，这要求我们探索崭新的内存解决方案。例如，Transformer模型的参数在短短两年内增长了240倍，而相对应的，单个GPU内存的容量却仅仅以每两年翻一番的节奏增长。AI模型训练对内存的需求往往是参数数量的数倍，因此对内存的需求迅速触及了现有技术上限，这里的“内存上限”不但涉及容量问题，还包括数据传输带宽的制约。

此外，数据通信已成为另一个影响计算力的瓶颈。这个问题涵盖了芯片内部、不同芯片之间以及AI加速器之间的所有通讯环节。在过去二十年间，尽管我们见证了算力的显著提升，存储技术从DDR一直进化到GDDR6x，并且接口标准也从PCIe 1.0a更新到NVLink 3.0，通信带宽的增长率却仅有30倍。从长远来看，仅依靠增加显存的方式来解决是不可行的。因此，我们需要更高效的方法来处理数据通讯问题，以便真正打破AI训练方面的限制。

存算一体技术彻底革新了解决存储墙问题的方案，优化了对Transformer架构算法的兼容性。在传统的冯·诺依曼体系结构中，数据处理过程依赖数据在内存和计算单元之间的频繁移动，由此产生了有限带宽导致的数据交换速度瓶颈，同时引发了存储墙和功耗墙问题。实际计算能力因此常常低于理论峰值。

通过在存储器中直接处理数据，存内计算技术有助于提高实际算力，可以突破因带宽限制产生的计算瓶颈。该技术利用存储器本身进行数据的直接处理，实现了存储与计算的同步进行，解决了传统架构下数据传输的限制，并降低了能耗。

特别是在人工智能和深度学习领域，这种计算理念优势显著。深度学习算法处理密集的矩阵乘法运算——一种乘累积过程，通过存算一体架构直接在存储结构内完成这些运算，既提升了能效比，也降低了延迟。生成式AI常用的Transformer结构，特别适合通过存内计算得以高效运作。

综合来看，从云服务、大模型一体化设备到智能计算中心和服务器，再到计算芯片供应商，这些构成了当前AIGC领域计算力发展的核心。这类技术的发展，对于推动人工智能领域的创新和应用至关重要。

(3) 中国人工智能开源生态的代表性组织和项目

随着人工智能大模型产业的飞速发展，模型的开放性成了业界热议的话题。基于模型源代码和训练数据公开程度的不同，大模型可被归类为开源和非开源两种模式，双方之间的争论日趋激烈。开源模式以其能够激发创新和促进生态系统繁荣而备受推崇，而非开源模式则以其出色的安全性和严格的质量控制著称。

在国内市场，一些大模型公司选择了走开源的道路，但对企业来说，是否开源需要评估其自身的技术优势和发展战略。例如，阿里巴巴选择了开源模式，以其强大的算力资源支持和对生态系统繁荣的推动为考量；而华为则保持其大模型的非开源状态，以确保“AI for Industries”战略下用户数据的安全性。

开源模式的优缺点分析如下：

优点：

- 创新发展：吸引更多开发者参与模型的改进，推动技术创新；
- 透明度高：代码和模型公开使得众多参与者能够进行直接审查，增加了安全信赖；
- 安全可靠：开源将使AI具有更大潜力来提升其创造力和协同能力，而对AI发展尤为关键的安全更离不开开源，AI技术发展的前沿是安全问题；
- 生态共建：便于建立积极的社群环境，促进大模型的持续升级与完善。

劣势：

- 滥用风险：可能会导致模型被不当修改或用于网络攻击等安全问题；
- 安全标准缺失：由于开源项目来自不同实体，其安全性和维护水平参差不齐。

非开源模式的优缺点分析如下：

优点：

- 质量管控：更好地维持模型的稳定性和质量控制，确保用户体验的一致性；
- 数据隐私：增强了用户数据的保护，减少了数据滥用的风险。

劣势：

- 创新局限：可能对技术多样性和突破性创新构成障碍；
- 垄断风险：增加了行业进入门槛，对新进入者造成排挤，有可能导致市场垄断。

在决策时，企业应选择适合自己的模式，并考虑到行业格局和长远发展。

阿里巴巴在自研大模型领域内采取开源行动，目标是通过开源策略使得广大中小企业和AI开发者得以更便捷、更及时地接触并利用其通用型人工智能模型“通义千问”。截至2024年2月，阿里巴巴已经分多阶段开放了多个版本的模型。

自2023年8月以来的开源历程概述如下：

- 2023年8月，发布了70亿参数的“通义千问”模型 Qwen-7B。
- 2023年9月，开放了140亿参数的“通义千问”模型 Owen-14B 以及其对话型模型 Qwen-14B-Chat。
- 2023年12月，对720亿参数模型 Qwen-72B、18亿参数模型 Qwen-1.8B 和专为音频设计的 Qwen-Audio 模型开放了源代码。
- 2024年2月，推出了“通义千问”开源模型的1.5版本，包括0.5B、1.8B、4B、7B、14B、72B六种规模的Base和Chat模型。

智谱AI:

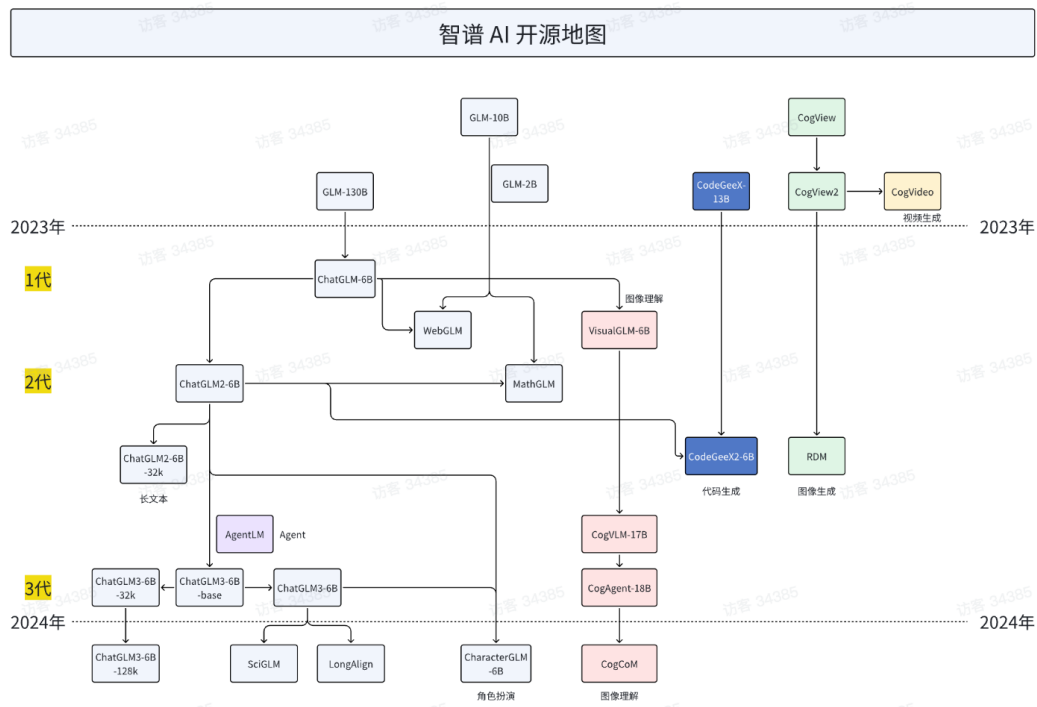


图 3 智谱AI开源地图

自2021年开始，智谱AI就致力于推进大模型的开源事业，陆续推出了包括GLM-10B、GLM-130B和CodeGeeX等一系列开源项目。迄今为止，智谱AI的开源模型数量已超过20个。特别是ChatGLM系列，在GitHub上获得了6万stars、7,300多个forks的卓越表现，在Hugging Face平台上下载量超过1300万次，连续四周位居趋势排行榜首。在科技部发布的《中国人工智能大模型地图研究报告》中，ChatGLM-6B模型的开源影响力位列首位（如图4）。此外，项目团队也被评为全球最高点赞AI机构之一，位列第五，并且是国内唯一入选的组织。

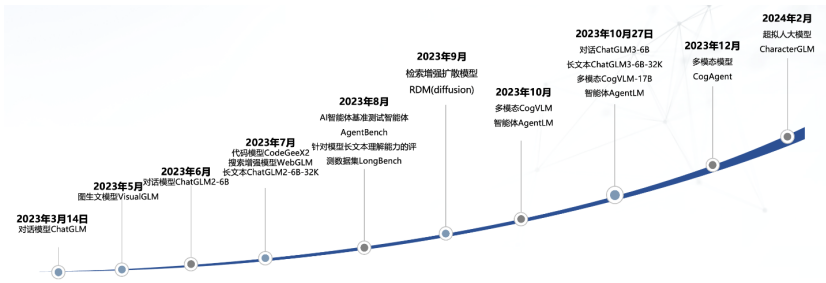


图4

3.3.3 中国人工智能产业发展现状与挑战

(1) 人工智能技术在中国的应用现状及典型案例

根据北京市经信局统计，截至2024 年3 月25 日，全国大模型企业总量达到了352个，北京大模型企业数量持续超过半数。大模型广泛应用于交通安全、智能制造、金融服务、教育、医疗健康等领域。百度、阿里、腾讯等在内的大型科技公司以及众多初创企业和科研院校，均在积极探索和实践人工智能大模型技术，研发和应用领域涵盖自动驾驶、语音助手、推荐系统、搜索引擎优化等。人工智能技术在中国的应用现状表现出以下5个显著特点：

- 一是产业规模持续扩大；
- 二是核心技术突破迅速，机器学习、深度学习、计算机视觉、自然语言处理等方面取得了重要进展；
- 三是行业应用深入。包括但不限于金融、医疗、教育、农业、交通、智能制造、智慧城市等。比如，在金融领域实现风险评估、反欺诈；在医疗健康领域辅助诊断、精准医疗；在教育领域个性化教学；在城市管理中实现智能化监控和公共服务智能化升级；

- 四是政策支持力度大。中国政府高度重视人工智能产业发展，将其列入国家战略层面，出台了一系列政策和规划予以支持。各地政府也在积极建设人工智能产业园区，吸引和培育相关企业，推动产业链协同创新发展；

- 五是人才需求旺盛。中国企业在人工智能人才的需求上极为迫切，许多企业愿意提供更高的薪资以吸引和留住具备AI技能的员工。同时，教育培训体系也在逐步完善，以满足市场对人工智能专业人才的需求。

当前应用案例主要集中在toC服务、生产辅助、硬件融合、辅助决策、智能制造5个方面。

2023年6月，修图AI应用妙鸭相机爆火。妙鸭利用AI算法优化照片质量，包括但不限于提升画质、修复瑕疵，并可以生成与用户上传照片类似但风格各异的新照片。用户上传自己的照片后，软件通过AI技术构建人脸模型并与其他模型合成，生成不同风格的写真照片，大大简化了传统修图流程。

智谱CodeGeeX是由北京智谱华章科技有限公司研发的一款基于大模型的智能编程助手，旨在通过先进的人工智能技术来提升软件开发的效率和质量。CodeGeeX结合了大规模预训练模型和深度学习技术，提供实时代码建议、自动补全、从自然语言生成代码、代码分析与改进、自动生成注释、搜索相关代码示例及跨语言转换等功能。

阿里云通义千问与联想合作推出的AIPC（即AI个人电脑）产品，该产品集成了阿里云的通义千问大模型，实现了人工智能与联想个人电脑的深度融合。该产品支持自然语言交互，提供便捷的信息查询、文件管理和智能办公等功能，利用通义千问的语言处理能力增强用户体验和设备智能化。

（2）人工智能相关企业和机构的发展状况

中国人工智能明星企业源于初代“AI四小龙”——旷视科技、商汤科技、依图科技、云从科技，是中国人工智能领域的领军企业，特别是在计算机视觉领域享有较高知名度。它们在人脸识别、图像识别、无人驾驶等诸多细分领域拥有前沿技术和创新应用。但因为AI企业普遍面临高昂的研发投入，长期处于亏损状态。例如商汤科技虽有巨额研发投入，但这也导致了较大的财务负担，即使上市，也需面对投资者对其商业模式可持续性的质疑。AI技术转化为实际收益的能力面临考验。尽管拥有先进的技术，但在大规模商业化落地过程中，如何找到稳定的盈利模式和市场切入点是一大难题，很多AI企业高研发投入与实际收入不成比例。随着全球及中国资本市场对AI赛道的投资热度趋于理性，投资方更加关注企业的盈利能力及可持续发展能力，这对于尚在烧钱阶段的AI四小龙来说，融资难度增大，不利于其快速扩张和持久经营。

大模型五小虎月之暗面、智谱、百川、Minimax、零一万物发展迅速。百川智能大模型Baichuan-13B-Chat(V2)在SuperCLUE的2023年8月排行榜中夺得国产大模型的榜首位置，表明百川在自然语言处理领域的技术水平处于领先地位，并且在2024年持续加强研发投入和模型迭代，扩大市场影响力。Minimax获得了阿里集团的高额投资，成为“大模型五虎”之一。智谱AI作为大模型独角兽之一，其发展态势同样值得关注。随着国内大模型市场的快速成长，智谱有可能在2023年强化了自身在学术研究、产业应用以及与合作方的合作关系，推出新的研究成果或商用解决方案。零一万物被列入大模型独角兽名单，意味着它在大模型领域有一定的技术积累和市场份额。2024年，零一万物继续聚焦在自己的核心技术方向上，比如通用型大模型的研发，或者在某个垂直领域（如工业、医疗等）取得重要成果。月之暗面是成立于2023年4月17日的创新型科技企业，专注于通用人工智能。同年10月即推出了全球首款能处理20万汉字输入的智能助手产品——Kimi AI，显示出其在AI技术研发上的强劲实力 and 创新能力。

（3）中国政府在人工智能领域的政策支持和规划

2018年10月31日，习近平总书记在十九届中央政治局第九次集体学习时讲话提出，加快发展新一代人工智能是我们赢得全球科技竞争主动权的重要战略抓手，是推动我国科技跨越发展、产业优化升级、生产力整体跃升的重要战略资源。2023年4月28日、5月5日中央会议关于人工智能发展提出三个第一次：第一次提出“通用人工智能”，第一次提出“产业智能化”，第一次提出“把握人工智能等新技术革命浪潮”。中国的各级地方政府积极响应国家对人工智能产业的战略布局，密集出台了多项政策，促进人工智能技术的创新与发展

其中，北京市于2024年启动“人工智能+”行动，适度超前建设5G、算力等基础设施。工信部则从三个方面出台支持政策：一是加快构建以先进制造业为支撑的现代化产业体系。改造升级传统产业、巩固提升优势产业、培育壮大新兴产业。推动新一代信息技术、智能网联汽车、航空航天、生物制造等新兴产业健康有序发展，加快北斗产业发展和规模应用。同时，超前布局建设未来产业。二是着力提升产业科技创新能力。加快推动以大模型为代表的人工智能赋能制造业发展。今年将开展“人工智能+”行动，促进人工智能与实体经济深度融合，推动人工智能赋能新型工业化。持续优化创新平台网络。三是推进信息化和工业化深度融合。当前，5G应用在工业领域深入推广，将适度超前建设5G、算力等基础设施，推动工业互联网规模化应用。同时，持续推进制造业数字化转型，积极建设智能工厂，实施中小企业数字化赋能专项行动，加快数字技术赋能，促进制造业向数字化、网络化、智能化发展。

总结来看，国内人工智能产业政策主要聚焦以下八个方面：

一是为产业提供政策扶持和资金支持，设立专项基金，为人工智能企业提供研发补贴、税收优惠、贷款贴息等财政支持。鼓励社会资本投入人工智能领域，设立引导基金吸引风投和私募基金加大对AI初创公司的投资力度。

二是加大力度支持基础设施建设如数据中心建设，为AI企业提供高性能计算资源和数据存储服务，并加强5G、物联网、云计算等新一代信息技术设施配套，为AI应用提供优质的网络环境。

三是支持技术研发和成果转化。扶持高校、科研机构和企业开展人工智能核心技术攻关，重点支持深度学习、自主无人系统、自然语言处理等关键技术的研发。

四是鼓励产学研用深度融合，建立科技创新联盟，促进科技成果转化成为生产力。

五是人工智能人才引进和培养，实施人工智能高端人才引进计划，提供优厚待遇吸引海外人才回国创业就业。支持企业与高校共建人工智能学院和研究生培养基地，加大人才培养力度，完善人才培养链条。

六是开放应用场景。支持人工智能在智慧城市、智慧医疗、智能制造、无人驾驶等领域的应用示范工程。出台场景清单，鼓励企业申报和实施人工智能重大应用场景项目。

七是标准体系建设和安全保障，包括加快人工智能技术标准和伦理准则的制定，推动形成完善的行业标准体系。强化人工智能安全技术研发，建立健全数据安全和隐私保护制度。

八是加强产业集群和生态建设。包括支持开源社区、平台发展中，设立人工智能产业园区，构建产业链上下游协同发展的一体化创新生态。举办各类人工智能峰会、竞赛等活动，搭建交流展示平台，推动区域间交流合作。

(4) 中国人工智能产业发展的瓶颈和制约因素

我们可从中美人工智能发展差距理解中国人工智能产业发展的瓶颈和制约因素。

美国依托AI基础理论、云计算、芯片、开源生态、人才、数据、资本等优势，在这一轮AI竞争中取得先发优势并持续巩固提升。

算法层面：我国基础大模型能力尚未达到GPT-4。国内基础模型在复杂推理、数学、代码等方面存在明显短板。美即将发布新一代基础模型，性能有数量级的提升，如OpenAI今年可能推出GPT-5，

性能将比GPT-4提升50%-100%。GPT-5及后续版本对算力、数据、资金、人才、工程化能力要求极高，赶超GPT-5及后续版本的难度会指数级提升，中国具备跟随能力的企业只有2-3家。尖端人才是大模型创新发展的重要基础，Sora团队仅13人，展现了年轻一代在AI大模型发展中的决定性作用。当今顶级的AI研究员工作所在地，美国占了60%，中国只有11%。

数据方面：我国缺少高质量的中文数据集。国内外基础大模型参数数据集主要以英文为主，如Common Crawl、BooksCorpus、Wikipedia、ROOT等，最流行的Common Crawl中文数据只占据4.8%。已上线中文数据集的整体数量和质量，与英文数据相差巨大，相当一部分内容非常陈旧。

算力层面：中国大模型企业需要迅速获得大量低成本、高效能的算力，而非芯片本身。单纯的芯片国产化替代暂时无法解决当前我国大模型面临的实际问题。美国已形成“芯片+云厂商”输出算力模式，芯片通过微软、谷歌、亚马逊等云厂商的公共云，形成低成本高效能的算力，服务大模型企业。但目前各地一哄而上建设“小散多”智算中心，不以公共云模式提供算力服务，导致算力利用率极低。

中美AI应用及产业生态差距拉大。基于应用场景的商业化闭环是AI竞争的焦点，是AI可持续发展的动能。当前美国大模型产业应用的速度、广度和深度都远超我国。一是美国AI产业应用速度快。OpenAI正式推出GPTStore，推动美国大模型进入了大规模的商业化应用阶段；二是AI大模型在美国的制造、法律、医疗等千行百业均有应用，微软、Adobe等美国科技企业正在持续使用大模型提升核心竞争力。近期微软市值更是超越苹果，成为全球市值最高的公司。美国AI原生企业数量全球领先，呈现出“千亿级+百亿级+十亿级+几亿级”不同估值规模、不同行业的AI原生企业，产业生态持续繁荣。

3.3.4 中国人工智能产业发展建议

(1) 加强人工智能大模型时代人才培养

大模型技术增强了人工智能的泛用性，它使得开发者们可以以更低的成本和门槛，针对特定场景研发出更优质的AI模型，从而加速应用型人工智能的落地。大模型被视作贯穿人工智能应用实施的关键路径，它触及的技术领域之广、知识和技能的交融程度之深，导致相关技术岗位既有高度的共性也有交集。在大模型时代，人才培育应着重考虑以下几个方面：

- 国家正加紧推进人才培养工程，现阶段，诸如职业认证和技能标准的定义正在制定中，同时，人工

智能大模型技术人才的培养模式正在创新之中，以期满足行业快速发展的需求和市场的期待。

- 需要构建并完善人工智能人才培养体系，致力于培育一支知识渊博、素质卓越、执行高效的人才梯队，为大模型创新投入坚实的人力资源。

- 需要建立一个多层次的人工智能大模型产业人才梯队，加强人才队伍的整体建设和层次分明的培育路径。

总体而言，塑造一支与大模型技术发展相适应的高质量人才库是当务之急，这需要国家和企业各级的全面推动，为大模型技术的创新与进步打造坚实的人才保障。

随着数字经济的蓬勃发展和工业数字化转型的不断深化，新技术正日益成为推动社会进步的关键动力。在这股趋势中，新兴行业迅猛增长，诞生了大量新兴职业类型，如人工智能工程师、虚拟现实工程师、物联网工程师、大数据分析师、工业机器人操控师、无人机操作员、智能制造工程师、区块链工程师等。

中国人力资源和社会保障部在《中华人民共和国职业分类大典（2022年版）》中初次列举了97个与数字产业相关的职业，这其中就包括了人工智能工程技术人员与人工智能训练师，正体现着这些职业的重要性和规范化趋势。

成为一名合格的人工智能工程技术人员，需适应国家职业技能标准，这不仅是技术能力和知识的展示，更是职业教育和人才培养中的关键环节，确保从业者具有标准化的技能，满足不同行业对人工智能人才的需求，从而推动整个行业的全面发展。

人工智能工程技术人员职业技能标准涵盖五个主要模块：

- 基本技能模块：要求从业者掌握数学、物理和计算机科学等基础学科，以及编程、算法和数据结构等专业技能，并应具备良好的语言沟通与信息检索能力。

- 应用技能模块：从业者需要精通自然语言处理、机器学习、数据挖掘和计算机视觉等专业领域知识，能够自如开发和实施AI应用。

- 管理和组织技能模块：要求具备项目管理、团队合作和沟通协调能力，确保项目进度高效率推进。

- 创新技能模块：激发创新精神，能够主导科研活动，用创新思维解决实际工作中的问题。

- 法律与伦理模块：强调法律意识和职业道德，确保在符合法规的框架下进行人工智能技术的研发

和应用。

这一职业技能标准构成了一个完备且系统化的框架，有助于人工智能工程技术人员提升专业能力，达到行业要求，推动人工智能技术的广泛普及与应用。

人工智能训练师的国家职业技能标准专门为那些运用智能训练软件的专业人士而设计，这些人员负责管理数据库、调整算法参数、设计人机交互、跟踪性能测试以及从事其它相关辅助工作。这一标准将职业技能分为五个等级，从L5到L1，涵盖数据采集处理、数据标注、智能系统的运行与维护、业务分析以及智能系统的训练与设计等多个层面，并为每个等级的职业能力设定了具体的要求和标准。

人工智能训练师资格考试由理论知识考试、实操技能考核和综合评审组成。理论知识考试主要以书面或电脑化考试的形式进行，考查候选人对专业基础知识的掌握程度；技能考核则通过实际操作或模拟演练来评估其专业技能；而综合评审通常针对技术层级较高的职称，如技师或高级技师，通过审核相关材料及口试来进行全面评定。

在理论考试部分，根据不同等级的资格认证，所测试的理论知识领域也会有所不同。五级和四级主要考察数据处理等基础知识，而从三级及以上则需要涉及更加专业的技能，如智能系统设计以及指导培训等。

国家不仅在职业资格证书中增添了与数字技术相关的职业，还在2019年推出了“专业技术人才知识更新工程实施方案”，以应对经济社会高质量发展需求。该更新工程在诸如新一代信息技术、生物技术、新能源、新材料等战略性新兴产业领域进行了大规模的知识更新与继续教育。

重点项目中，数字技术工程师的培养项目专注于智能制造、大数据、区块链和集成电路等数字技术新兴职业。凭借技术创新和数据赋能，以及遵循国家职业标准和新职业培训教程，该工程对从业人员进行规范化培训和社会化评价，计划每年培训约8万名数字技术相关技能人员。

截至2023年4月，人社部在新闻发布会上宣布，这个项目已经初步构建了包括政策、组织、标准、培训和评价在内的完整体系，并计划拓展项目范围。北京市在7月提出的培训实施方案计划每年培养一万名数字技术技能人才，其他如重庆、山东、湖南等省市亦制定了相应的培养计划。

对于人工智能人才的供应和需求状况以及人才培养现况的分析，我们通过对人力资源社会保障部针对新兴职业的就业市场前景报告进行了综合梳理。在未来五年内，预计对与人工智能、数字化以及智能化紧密关联的新兴职业需求将大幅增长。

随着人工智能领域的蓬勃发展，高校作为培养相关人才的重要基地，在过去三年里极大地增加了人工智能专业的开设。根据《2022年中国人工智能人才发展报告》，已有440所本科高校分四批次获批设立该专业，占全国1270所本科高校的34.6%。自2019年以来，人工智能专业在教育部的支持下不断扩张，首批由35所高校开设，紧接着在2020年、2021年和2022年分别有180所、130所和95所高校获得批准。在这些被授权的440所院校中，包括81所985和211工程院校，113所一本院校，以及24所二本院校。这些人工智能学专业遍及全国29个省区市。

与此同时，众多企业也正在加强与高校们以及行业组织的合作，提升人才培养的力度。国内的头部企业如华为、阿里巴巴、腾讯和百度等均与高校合作进行产教融合，通过建立产业学院、共建专业、定制培养等多种方式进行人才培养，同时也利用企业的平台和资源开展人工智能人才的培育工作。

从国家层面、教育机构到企业，对人工智能人才的培育日益给予重视和投资。然而，我国在人工智能技能人才方面目前存在以下三大缺口：

- 致力于推进前沿技术和基础理论研究的顶尖人才：这一群体以其在人工智能领域的深厚学术背景及研究资历，对技术前沿的持续推动和理论创新至关重要。由于这类人才较为稀缺，处于探索阶段和前沿的大型企业愿意支付高成本以吸纳他们。企业倾向于向国内外顶尖高校挑选充满潜力的毕业生，提供优质的研究资源和项目，从而进行人才孵化。
- 能够将理论与实用相结合的算法模型设计人员：他们擅长将尖端理论与实际的算法模型互补，并在算法设计和优化上拥有丰富的实践经验。他们将自己的专业技能投入到设计高性能的算法模型中，以提升人工智能系统的执行效率。大型企业，特别是在人工智能领域的引领者，争相招募这类人才以推动公司的研发和创新。
- 能够将人工智能技术与行业需求相结合的应用型人才：这些人才紧密结合行业需求，深刻理解各种应用场景，并利用人工智能技术解决实际问题，降低成本、效率以及提升产品和服务的质量。这类人才因应用范围广泛而需求量较大，企业在招聘时会特别关注其行业背景和实践经验，以确保能与公司的具体需求相融合。

为适应人工智能这一新技术革命不断涌现的人才需求，政府与教育机构亦相继出台了系列措施以培育和引导人才发展。2023年4月28日，中共中央政治局会议强调了重视通用人工智能发展的重要性，并提到要建立相关的创新生态和风险防范机制。网信办发布的《生成式人工智能管理服务暂行办法》旨在为AI大模型的长远发展建立规范。北京市进一步出台了《北京市促进通用人工智能创新发展

的若干措施（2023-2025年）（征求意见稿）》，专注于加强算力资源供给、高质量数据提供、探索大模型技术体系以及推动技术创新等方向，提出了实施21条具体政策和措施。

在国家的整体战略和政策指引下，配合中国软件行业协会的服务目标，下面是我们对人工智能大模型技术人才的培养提出了以下思路和建议：

- 结合“十四五”规划和行业需求，建立与本行业及目标受众适宜的人才培养目标、体系和服务模式。打造连接企业、行业 and 教育的“桥梁式”人工智能大模型技术人才培养体系。
- 贴近行业目标和企业需求，以工作实际所需的技能和知识为基础设计培养与认证体系，加强企业参与并借助专业教育资源，实现产教深度融合。
- 利用行业研究资源和区域发展需求，针对性地提供人才培养，与当地服务建立人才实训基地，进行专业技能培训和认证，满足企业对人才的迫切需求。

（2）促进产业应用融合创新

ChatGPT的发布预示了大模型技术时代的到来。虽然人工智能技术已实现显著突破，但其产业化和商业化的仍面临不少挑战。现阶段，如何实现大模型的有效落地并使其赋能各行各业，已成为业界关注的中心议题。对大模型产业生态进行生态学层面的分解，可以将其分为四个层级：数字基础设施层（L1）、通用大模型层（L2）、行业模型层（L3）以及场景模型层（L4）。构建一个生态型的商业模式至关重要，需推动各生态参与者之间的能力共享与组合，灵活匹配核心能力和特有能力和，以此推进大模型在各行各业的应用，共创未来的共享经济。

人工智能技术从早期的探索到现在已经经历三个发展阶段，但在商业和产业应用方面始终难以突破。1956年的达特茅斯会议首次引入了“人工智能”术语，当时人工智能领域主要局限于理论研究。1970年代，计算能力和算法取得进展，但数据量的限制阻碍了进一步发展。21世纪第二个十年，硬件如GPU的提升、算法优化以及大数据集的可用性为人工智能开启了新纪元。神经网络的演进以及深度学习的兴起代表了AI领域的一个转折点。深度学习训练的复杂网络推动了大模型的发展，并成为现代人工智能的关键支撑。但至今，人工智能在商业化方面仍未实现大规模应用，主要困境是寻找适用场景与行业知识的不足。

大模型引领了AI的新纪元，迁移学习是支撑大模型的关键技术之一。Transformer架构的提出和模型如BERT、GPT的成功应用标志了自然语言处理领域的重大突破。大规模图像数据集如ImageNet

为图像处理大模型奠定了基础。现代大模型如GPT系列在不同领域崭露头角，人工智能生成的内容（AIGC）引起了广泛关注，展现出大模型的应用潜力。

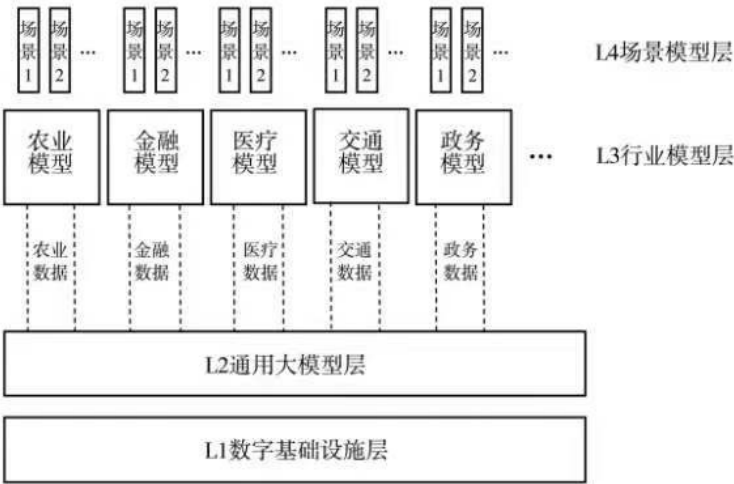
大模型需要基于生态视角推进产业化发展。本研究将分析大模型产业生态的四个层次，并讨论如何构建生态型商业模式，将大模型的力量与各行各业的实际应用相结合。为了进一步加强大模型和行业之间的紧密联系，本文提倡应用生态培育的理念，与多方生态伙伴合作，集合创新力量，共同构筑大模型产业生态系统。

本研究集中于以下四个生态层次，分析其特征及作用：

- L1 数字基础设施层：提供必要的基础设施以支持大模型训练和运行。
- L2 通用大模型层：作为广泛应用的基础模块服务各行业。
- L3 行业模型层：结合特定行业需求的定制模块。
- L4 场景模型层：更细致地针对特定场景或问题提供解决方案。

概述大模型产业化和商业化应用的发展潜力，并探索在现代商业生态中大模型如何赋能各行各业，推动数字化转型和智能化升级。

本研究提出，大模型产业生态结构可细分为四个互相支撑的层次，包括数字基础设施层（L1）、通用大模型层（L2）、行业模型层（L3）以及场景模型层（L4）。



大模型产业生态的生态层次图

- L1 层次 — 数字基础设施层：数字基础设施构成了大模型产业发展的根基，涵盖硬件如CPU、GPU、专用智能处理器，软件包括数据库、中间件以及操作系统，云计算服务，以及数字网络设施例如5G和工业互联网。这个层次提供了处理大量数据和运行复杂模型所需的关键算力支持，云服务为大模型带来灵活的计算资源，而高性能网络保证快速数据输。
- L2 层次 — 通用大模型层：通用大模型提供跨行业普适的AI能力，如自然语言处理、图像识别和语音处理等。它们需要处理海量数据以提炼出通用能力，并通过不断迭代和优化以跟上技术进步。以GPT系列模型为例，其模型规模和能力持续扩展，例如，GPT-4引入多模态处理的能力，能够处理包含视觉信息的复杂任务。
- L3 层次 — 行业模型层：该层次基于通用模型进一步定制化，结合特定行业的数据和特征，打造出针对特定行业需求的大模型。这种行业专属模型从行业数据中提取关键特征，解决行业独有的挑战，例如健康诊断、风险评估与优化建议等。
- L4 层次 — 场景模型层：场景模型层更进一步在行业模型的基础上进行细分，针对特定行业内特定场景或问题提供解决方案。这些模型针对确切的问题定制设计和训练，并融合特定场景的数据和专业算法进行构建，如自动驾驶交通场景下的数据处理与决策支持。

大模型产业生态的这种层级化构造旨在确保从基础设施到应用层面的有效衔接，使不同类型的大模型能够适应并解决不同层级的问题。如自动驾驶的智能交通应用，场景模型层可能专注于优化载货路径以减少运输成本，或者在Robotaxi场景中确保严格的安全性和适应性。这种分层的生态结构有助于将AI和大模型的威力更准确地配对到不同行业和场景的具体需求。通过这种方式，企业和组织能够基于通用模型能力，以更加高效、精准的方式推动领域内的创新和问题解决。

矿山行业的大模型应用也呈现多样化的场景需求。以华为盘古矿山大模型为例，已涵盖了从采煤、掘进、到物资运输等多达21个不同场景的应用。每个场景下，AI大模型的焦点都因需求而异。例如，在矿山主运输系统中，大模型承担起“高速公路”般的角色，其AI能力主要集中在智能监测与识别，能够取代传统的人工巡检，以全天候的监控实时识别异常，为系统安全和高效运行提供保障，并且减少人工成本和风险。在掘进环节，由于涉及到的操作规范对于作业人员的安全尤为关键，大模型的主要任务是执行精确监测，确保作业流程的标准化执行，从而最大化地维护作业人员安全。

整个大模型产业生态由以上提及的四个层次构成，其中任何一个层次的缺失都将影响大模型产业的整体竞争力。全球各国亦在大模型产业生态的构建中争夺先机，以发布数量为例，美国和中国目前处

于领先地位。截至2023年7月，美国累计发布了114个大模型，而中国的发布数量也实现了迅猛增长，达到130个。不过，中国在大模型生态的发展状态依旧是“量多质未优”，在理论、算法和技术路径上仍依赖国外。在L1数字基础设施层和L2通用大模型层，美国设置了大模型发展的主要技术框架，而在中国，尽管如百度、华为和商汤等企业进行了基础框架的自研工作，但国产大模型有不少是基于国外开源模型进一步训练和微调的结果。L3行业模型层和L4场景模型层方面，国外以谷歌、微软、Meta为中心的大模型生态系统更为成熟，企业通过API接口来调用大模型，形成了一套完善的生态体系；中国则呈现多点开花的竞争局面，多个行业正探索大模型的应用潜力。

展望未来，中国在构建大模型产业生态上需要发挥自身产业特色，积极创新和提高在L1和L2层次上的自主能力，同时利用本国产业基础和市场优势，精准施策于L3和L4层次，以促进大模型在更多行业的应用，并推动行业智能化转型升级。

大模型产业生态通过其多层次结构彰显了生态型商业模式（EMB）的核心思想。EMB是由多种多样的参与者联合构成的综合商业模式，形成一个不断演变的生态系统。与传统商业模式相比，EMB不再局限于单一产品或服务的价值创造和价值获取，而是涉及多样化合作方的共同发展。目前的商业模式研究主要集中在战略和创新两个领域，强调公司能力和对外部创新的响应。EMB概念重塑了这一研究范式，突出了参与各方在价值共创中的互动与合作，并增强了共同创新的重要性。

在大模型产业生态的构建中，通用平台和生态伙伴的共存是实现生态型商业模式的基础。例如，L1数字基础设施层和L2通用大模型层形成了广泛对外开放的平台，这些平台作为技术供应商，提供必要的支持和服务，勾画出了广阔的应用蓝图。而各类行业的生态伙伴，则利用这些平台下的资源和服务，根据自身特定的需要，发展出适合自己行业和场景的定制化解决方案。

大模型产业的发展需要众多生态参与者的相互配合与支持，包括政府、行业协会、数字企业、行业巨头和中小企业等。政府应该构建良好的数字基础设施，引导和规范参与者间的交互与合作，促进产业的规范化发展。行业协会和企业则应该联手制定研发计划和相关标准，推动标准化和创新合作，尤其是行业领导者发挥示范作用至关重要。同时，必须具备开放的生态融合思维，使各个领域的参与者在在大模型的支撑下开展创新实践，拓展新的应用场景，实现共赢。

生态型商业模式突出了对能力共享和组合的需求。在这种商业模式下，企业不再仅提供单一产品服务，而是要充分展现自身特有的技能，并与生态伙伴的能力相结合，以形成适应多变场景需求的解决方案。通用平台扮演着提供核心AI能力的角色，如云计算、AI模块等。同时，为了发挥大模型在各细

分行业的潜力，还需融合各行业伙伴的专有知识和技术。

通过这种跨领域合作和能力整合，生态型商业模式能推动动态、互补的发展体系。利用通用平台核心能力的伙伴，可以在较低的投入成本下开展创新和适应市场需求，而成熟的平台企业则能借助于这些多元化合作丰富自身的行业洞察和解决方案集。这不仅实现了资源的优化配置，而且为打造应用广泛的共享经济铺平了道路。

借助生态型商业模式推动大模型产业生态的成熟，将为多个行业的数字化转型注入助力。通用大模型提供的核心功能构成数字化转型的基础，支持企业进行高效的数据分析和决策，提升生产和服务效率。这促使企业不断发展自身的特色能力，加快构建全面的大模型技能生态。大模型的兴起预示着工业和消费互联网领域更深层次的融合，如在工业制造中实现生产过程的智能监控与优化，在消费服务中提供更加精准的个性化推荐，这些都有助于实现全面的数字化生态和智能服务。

然而，面临大模型产业生态建设，我们还需解决若干挑战。首先是在标准制定与市场监管方面，必须协同行业伙伴共同推动技术标准、产业规范和市场标准的建立，确保技术互通和市场的良性竞争。其次是数据市场的构建和管理，此项对于大模型的训练至关重要，加强数据流通平台的构建并妥善处理数据隐私和安全，是推动大模型生态进步的关键。此外，知识产权是大模型发展的核心问题之一，必须确立明确的产权归属、使用许可以及成果分享的规则，保证生态中各方的合理利益。

展望未来，大模型时代的到来将激励各行各业发挥其潜力，从而构建一个广泛合作与创新的生态系统，引领科技革新与经济增长。随着数字化转型的推进，大模型还将为全球面临的环境、能源和医疗等挑战开辟新的解决方案。在这新时代中，跨领域的合作、制度完善和生态发展将是实现大模型商业化落地和创新的关键，为全球共同繁荣提供强大的动力。

（3）完善政策法规和伦理治理

我国大模型的发展在法律和政策层面面临以下几个关键问题：

数据安全与隐私保护：随着大模型对海量数据的依赖性增强，如何确保数据收集、存储和使用的合规性，避免侵犯公民隐私和信息安全成为一个重大挑战。解决办法可能包括制定和完善相关法规，比如严格执行《网络安全法》《数据安全法》《个人信息保护法》，要求企业在使用数据前获得合法授权，并采取必要的脱敏、加密等措施保障数据安全。

- **数据产权与使用权界定：**由于大模型训练所需的大量数据往往来源于多个主体，数据产权不明晰

可能导致法律纠纷。建立清晰的数据权属制度，明确数据来源和使用权分配，鼓励数据合法合规流通和共享。

- **法律责任认定：**在大模型输出结果产生错误、误导或侵权时，如何准确界定法律责任归属也是一个难题。需要进一步细化相关法律法规，确立合理的过错责任原则和追溯机制。
- **技术出口管制：**鉴于大模型技术的战略意义，国家可能会加强对其出口和技术转移的监管，制定相应的出口管制清单和技术审查程序，确保核心技术不被滥用或流向不当。
- **行业标准与监管滞后：**大模型技术发展迅速，而对应的行业标准和监管机制可能存在滞后现象。政府应当引导和推动建立适应新技术特点的标准体系，同时强化市场监管能力，对违规行为实施有效惩处。

为解决目前人工智能治理面临的立法、标准、监管问题，应加强立法工作，紧跟科技发展趋势，修订完善现有法律法规，填补法律空白。制定大模型领域具体的技术规范和行业标准，指导企业和研究机构合规发展。设立专门的数据监管机构，严格监管数据采集、处理和应用过程。提升公众对数据权益的认知，推动形成全社会共同维护数据安全的良好氛围。加强国际合作，借鉴国际先进经验，参与全球数据治理规则的制定，促进国内外市场的有序竞争和健康发展。

（4）推动开源生态建设，加强开放协作

中国加速发展模型开源社区，不仅有助于缩短技术研发周期、降低成本，还有利于激发全社会的创新活力，促进跨界融合和资源共享，培养大批高技能人才，并在全球人工智能竞赛中赢得战略主动权，从而推动整个产业的可持续、健康发展。

推动我国人工智能开源生态建设，发展开源社区，在政策引导上，可通过政策倾斜、设立专项基金等形式，鼓励和资助人工智能开源社区的建立和发展，如科技部在科技创新2030新一代人工智能重大项目中，就强调了开源、开放的原则，并引导国家项目支持研发成果的开放共享。在激励机制上，可设立开源项目奖励计划，表彰和资助优秀的开源项目团队和个人，如设立开源项目奖项，为开源贡献者提供荣誉与物质激励。启智社区、中国科学院牵头的“鹏城实验室”等，都是国家层面对开源社区建设的有力支持。同时支持企业开源技术，为主动开源的企业提供税收优惠与补贴，可以给予相应企业税收减免或财政补贴，减轻企业在开源技术研发过程中的经济负担。对于企业开源的技术成果，可以考虑在一定条件下豁免部分专利申请费用，同时，在政府采购中优先考虑采用国产开源AI产品的项目。

同时补贴开源应用落地。设立专项资金支持基于开源技术的人工智能应用示范项目，鼓励开源技术在各行业的实际应用，对成功的应用案例给予补贴或推广奖励。对于由开源技术孵化出的产品和服务，可通过创新券、科技成果转化基金等方式给予支持，降低商业化初期的成本和风险。由政府 and 产业园区通过举办开源创新大赛、提供创新创业支持等方式，推动开源AI技术在本地产业的应用。

最后，营造良好的开源文化和环境。一是在高校和职业培训中加入开源文化课程，培养具有开源精神的AI人才，同时加强对开源社区建设的理论研究和实践经验的传播。二是参与国际合作，积极参与全球开源社区，与国际领先的开源组织合作，引进国外先进的开源理念和技术，同时推动中国AI开源项目的国际化。

(5) 提出针对现状问题的解决对策建议

大力发展公共云，为我国大模型的训练和产业化应用提供优质高效低成本的算力。建议将大力发展公共云纳入国家“十五五”规划，加大中央和地方各级财政专项对公共云核心技术攻关及应用研究支持，政府和央国企数字化项目要优先采购公共云算力服务和产品，加强对“小散多”智算中心建设的约束，明确算力中心以公共云提供服务的比例要求，更好保障我国大模型的训练和规模化应用，并进一步提升已建算力中心的可持续运营能力。

坚持把大型科技平台作为中美AI竞争的主力军。在AI国家战略布局上，应把具备体系化竞争能力的大型科技企业作为中美AI竞争的主力军，在美实施非对称竞争和打压背景下，聚焦国家战略资源，以赛马形式大力支持大型科技企业加速研发下一代AI多模态大模型，全面提升AI芯片、云计算平台、基础模型、开源生态、场景应用等方面的体系化竞争能力。

坚持扩大公共数据开放，加快形成高质量的中文多模态数据集。加快构建统一的公共数据开放基础制度规范，扩大面向AI产业的公共数据供给范围，加大政务类、医疗类、数学类、科技类、广电媒体类、教材教辅类等多模态数据开放力度，推动各类图书馆、知网等专业数据库与AI企业合作，支持龙头企业牵头制定数据标注规则和数据标注质量评估体系，加快构建中文高质量数据集。

坚持开源开放，促进AI产业技术生态繁荣发展。加强AI开源生态的顶层设计，将AI开源生态建设纳入国家相关规划、抓好落地实施，探索利用科技专项等政策举措支持科技平台企业做大做强AI模型开源社区，推动在政府资助下形成的AI研发成果的开源，扩大AI开源供给。支持企业加速应用基础大模型、研发应用行业模型和企业专属模型，通过“用模型”反哺技术创新。

加大力度吸引和培养全球AI人才。支持民营企业AI人才参与国家重大科技专项、重点研发计划等顶层设计和执行。支持民营企业AI专家参与国家杰出青年科学基金、长江学者奖励计划、两院院士等高层次人才评选，并在政策扶持、经费支持等方面享有与国有企事业单位同等待遇。对于符合条件的高端AI人才，实行优先落户政策，提供税收优惠，保障子女就学、家人就医等，解决人才后顾之忧。

坚持审慎包容、分类分级的治理态度和思路，鼓励AI技术创新和产品研发。考虑到中美AI竞争的实际情况，对还处于发展初期的AI大模型技术和应用，建议结合其技术特征和发展规律，在监管模式上更注重事中、事后监管，对应用场景风险较低的新技术新应用（如电商、生活服务、企业数字化转型等），以行业自律等方式予以治理，留出更多试错、容错空间，保护企业投资意愿和创新活力。

3.4 RISC-V行业研究报告

3.4.1 RISC-V概述

RISC-V作为开源指令集，经过近几年的快速发展，已经成长为仅次于x86和ARM的国际第三大指令集架构。在2015年成立的RISC-V国际基金会，旨在为基于RISC-V ISA的软硬件开发者建立一个开放、协作的社区。截至2024年4月底，RISC-V国际基金会已接纳全球70多个国家的4100多个会员，会员单位相比2019年的435位，增长了超过942%，相比2022年底的数据又增长了超过125%。RISC-V国际基金会的会员可以参与RISC-V指令集规范以及相关软硬件生态的开发，并决定RISC-V未来的推广方向。

在RISC-V国际基金会理事会26名理事中，与中国相关的有8家，包括北京开源芯片研究院、阿里巴巴、RIOS实验室、希姆计算、晶心科技、飞腾、华为、成为资本；与美国相关的有11家，与欧洲相关的有6家，南美洲（巴西）1家。

在RISC-V国际基金会中包括高级会员（Premier Members）25家，其中中方成员12家，占比达到48%；战略会员（Strategic Members）174家，其中中方成员41家，占比达到24%；社区会员（Community Organization Members）177家，其中中方成员14家，占比达到8%。会员中包含科技公司、芯片设计公司、IP公司、软件开发企业、高校院所等多种开发主体。

近年来，RISC-V在国际市场上占据越来越重要的位置。据RISC-V国际基金会披露，截至2022年，市场上量产的RISC-V处理器已经达到100亿颗，其中有50%来自中国。根据市场研究公司Counterpoint

Research预测，到2025年，RISC-V处理器累计出货量将超过800亿颗，年复合增长率高达114.9%，届时RISC-V将占据全球14%的CPU市场、28%的IoT市场、12%的工业市场和10%的汽车市场。

目前，全球RISC-V产业生态已经初步建立。美国、欧盟、以色列、印度等国家的众多开发者纷纷加入到RISC-V生态中，不断推动RISC-V技术与产业发展。其中，美国国防高级研究计划局连续多年专项支持RISC-V研究和实用化，而谷歌、英特尔、英伟达、高通等软硬件巨头公司也加入RISC-V的阵营中；欧盟不仅在《欧盟芯片法案》中强调要发展RISC-V，还宣布投资2.7亿欧元来开展RISC-V小芯片项目；以色列为发展本土IoT与高性能计算芯片，国家创新局投入1360万美元建设RISC-V协作平台；印度政府将RISC-V定为国家事实指令集，推出数字印度计划。

中国也高度重视RISC-V的发展，有数据显示，中国有超过400家企业关注、参与和使用RISC-V指令集进行开发。北京、上海、深圳、武汉、杭州等城市陆续推出一系列科研、产业化支持政策，以加强本土基于RISC-V的芯片设计研发能力。同时，中科院、阿里巴巴等科研院所与企业也在持续推出基于RISC-V的芯片产品。目前我国已经形成了较完整的RISC-V产业链，产品覆盖消费电子、通讯、汽车、通用计算等领域，具备产业化优势。

3.4.2 中国的技术贡献

(1) MCU领域

● 国内企业积极推动RISC-V车载芯片发展

2023年8月，我国RISC-V芯片设计企业芯来科技宣布，芯来NA系列CPU IP NA900顺利获得了ISO 26262最高汽车功能安全等级ASIL D的产品认证证书，而NA900也成为全球首个获得ISO 26262 ASIL D认证的RISC-V CPU IP产品。

2023年7月，由东风公司牵头成立的湖北省车规级芯片产业技术创新联合体完成了国内首款基于RISC-V指令集架构车规级MCU芯片的首次流片。

此外，比亚迪也在与华为合作，双方在基于RISC-V架构的芯片上联合开展研究。奇瑞也在与国内芯片厂商合作，积极推动整个RISC-V的测试架构，例如基于RISC-V的客户端和RISC-V芯片的测试标准，来确保芯片的可靠性。

● RISC-V芯片在物联网安全领域应用越发广泛

2023年7月，由赛昉科技、微五科技与名气家三方联手研制的RISC-V物联网安全芯片“港华芯”销量正式突破一百万颗。“港华芯”采用赛昉科技与微五科技的RISC-V芯片硬件基础，以及名气家安全协议与平台规范，达到国密二级安全标准，支持SM1/2/4/9等国密和国际加密算法，主要应用落地场景是智能燃气表。

2023年3月，苏州国芯科技宣布其基于RISC-V指令集架构研发的CCM3310S-L、CCM3310S-LP安全芯片已规模化应用于eSIM智能穿戴、智能门锁安全、ETC OBE-SAM、燃气表安全SE等物联网应用领域，出货量已达数百万颗。该安全芯片系列均采用40nm EFlash工艺设计，支持DES、AES、RSA、ECC、SHA等国际算法，SM2、SM3、SM4等国密算法，安全等级达到《安全芯片密码检测准则》第二级要求。

● 推出基于RISC-V架构的AI技术框架MCU

2023年12月，华为海思推出了一款基于RISC-V内核并结合端侧嵌入式AI技术的A²MCU。A²MCU聚焦行业专用（Application Specific）和嵌入式AI技术，将AI领域超轻量级的技术框架、极致性能的推理要求、方便快速的部署能力与MCU深度融合，为MCU行业客户探索智能化应用提供新的选择。海思表示，A²是海思针对家电、能源、工业、汽车等领域推出的全新解决方案，它不仅涵盖了基于RISC-V的系列化的MCU，还包含了高性能兼容ARM指令集的MPU，以及与之紧密配合并优化的openEuler操作系统。

(2) 芯片系统领域

● 推出基于RISC-V的多芯片设计系统

2024年1月，中国科学院计算技术研究所推出了一种名为“Zhejiang”（浙江）的“大芯片”，即是一种直接使用整个晶圆制造的多芯片设计系统。该芯片采用Chiplet设计，共包括16个芯粒，而每个芯粒内有16个RISC-V核心，总计256核心，且均支持可编程、可重新配置。这颗“大芯片”基于22nm工艺制造，由超过1万亿个晶体管组成，总面积数千平方毫米，采用小芯片封装或计算和存储块的晶圆级集成。不同核心之间通过网络芯片（Network-on-Chip）、同步多处理器（SMP）实现互连，不同芯粒之间再通过D2D（Die-to-Die）接口、芯粒间网络接口（Inter-chiplet Network）实现互连，并使用了2.5D中介层行封装，以实现小芯片之间的内存共享。

● 发布利用AI全自动生成的CPU芯片

2023年6月底，中科院计算所等机构用AI技术设计出了世界上首个无人工干预、全自动生成的CPU芯片——启蒙1号。这颗完全由AI设计的32位RISC-V CPU，相比GPT-4目前所能设计的电路规模大4000倍，并可运行Linux操作系统，且性能堪比Intel 486。为了减少人力和资源投入，研究人员采用AI技术直接从测试用例的输入/输出（IO）自动生成CPU设计，无需工程师提供任何代码或自然语言描述。该方法在5小时内生成了超过4,000,000个逻辑门的32位RISC-V CPU，甚至自主地发现了包含控制器和运算器等在内的冯诺依曼架构，为后续发现新的体系结构优化知识提供了参考。

● 推出面向多领域应用的RISC-V处理器芯片

2023年11月，阿里平头哥发布三款基于RISC-V内核的玄铁系列处理器：玄铁C920、玄铁C907、玄铁R910。其中，玄铁C920升级支持最新的Vector1.0标准，可实现更精准、稳定地分配任务，较上代提升了高至3.8倍的AI性能，可用于Transformer模型，适合机器学习、自动驾驶等领域。玄铁C907首次实现了独立矩阵运算（Matrix）扩展，提高了计算密度和计算并行能力，充分挖掘出RISC-V架构的AI原生优势。玄铁R910同时支持Cache以及TCM存储架构，可以在保持高可靠的基础上提升系统实时性，可应用于存储控制、网络通信、自动驾驶等领域。基于软硬协同研发的这三款玄铁处理器，提升了加速计算能力、安全性及实时性，将加速推动RISC-V在自动驾驶、人工智能、企业级SSD、网络通信等场景和领域的商用落地。

2023年8月，智芯科正式发布基于精度无损SRAM存内计算（CIM）的超低功耗神经网路处理器芯片——AT700，可用于安防监控、智能家居、机器人、智能驾驶，智能后视镜CMS等领域。

● 推出首个适配RISC-V的云操作系统及虚拟机

2023年5月，中国电信研究院牵头，联合天翼云公司、中科院软件所PLCT实验室，完成CTyunOS在RISC-V指令集芯片的适配测试，共完成4100+基础软件包和1100+拓展包编译适配，并在国内合作伙伴首台RISC-V服务器板卡成功运行，推动CTyunOS成为目前业界首个适配RISC-V云服务器操作系统。同时，中国电信研究院基于CTyunOS及欧拉开源生态，还成功研发出业界首个支持RISC-V的云原生轻量级虚拟机TeleVM，并在RISC-V CPU上完成软硬件协同测试验证。据官方介绍，TeleVM基于Rust安全语言重写了虚拟机并进行RISC-V架构适配优化，结合RISC-V精简指令集特性，裁剪冗余功能，实现了应用安全隔离，同时降低资源消耗并提升响应速度，适用于可信云原生、Serverless、函数计算等新型弹性云计算业务场景。

● 发布首款RISC-V开发笔记本电脑

2023年8月，深度数智发布了全球首款RISC-V开发笔记本电脑ROMA，并默认搭载openKylin（开放麒麟）国产操作系统，可体验原生RISC-V开发及RISC-V软件生态系统。据悉，该电脑采用4核RISC-V12/22nm处理器，配有4/8/16GB内存，未来还将支持ChromeOS等。

（3）人工智能领域

● 推动RISC-V在AI部署方面加速落地

2023年世界互联网大会乌镇峰会上，希姆计算联合国鑫发布了基于RISC-V架构的大模型系列一体机，包括2U4卡、4U8卡机，支持部署6B、7B、13B参数规模的大模型，这也是专门针对私有化部署提供的系统级解决方案。希姆STCP920系列采用12nm工艺，是基于RISC-V的人工智能计算加速卡，具备较高的云端AI计算能效比和可编程性，混合支持FP16/INT8数据类型。

● 推出面向AI的RISC-V全栈软硬件平台

2023年RISC-V中国峰会上，阿里平头哥更新自研一站式AI部署套件HHB，实现典型网络性能比第三方工具平均提升88%，并增加支持运行Transformer、TensorFlow、PyTorch等170余个主流框架AI模型，让RISC-V真正成为AI算力的新选择。基于此，平头哥发布了首个面向多媒体AI增强场景的RISC-V全栈软硬件平台。该平台将RISC-V扩展性的新型Vector、Matrix及第三方硬件进行算力抽象，创新接入OpenCV与CSI-NN等弹性计算库，深度融合多媒体处理流程，形成面向业务的流水线设计，方便用户在流水线的不同步骤上进行AI增强优化，为检测、分类、跟踪和超分等各类应用提速。

3.4.3 中国RISC-V的发展实践

（1）举办全球性会议赛事推动RISC-V发展

● 成功举办RISC-V中国峰会

2023年8月23日至25日，2023 RISC-V中国峰会（RISC-V Summit China 2023）成功在北京举办。RISC-V峰会由RISC-V国际基金会组织，每年在北美、欧洲、中国分别举行，是全球RISC-V领域规格最高、影响力最广的年度峰会。本届峰会为第三届中国峰会，以“RISC-V生态共建”为主题，采用“主论坛+主题报告+展览展示+同期活动+Poster”的方式，涵盖了100余个主题报告、16场同期活动、18个企业展位、16个社区展位、67个Poster及31个Poster预约交流。来自RISC-V国际基金会、业

界专家、企业嘉宾及社区代表等专家学者，结合当下全球新形势，对RISC-V产业发展进行了深入探讨。

● 举办RISC-V应用创新大赛

2023年7月到11月，第三届玄铁杯全球RISC-V应用创新大赛成功举行。本次大赛采用内置玄铁RISC-V处理器的3款量产高性能开发板，基于平头哥全新开源的安卓13系统SDK等软件栈，开发者可首次在RISC-V高性能硬件上体验安卓、openKylin等操作系统并直接开发相关应用。经过四个月的报名及开发阶段，本次大赛共计2012位开发者参赛，收到511份创意方案，参赛作品覆盖Vector加速、YOLO、机器人、车载设备、可穿戴设备、工业设备等领域。

（2）积极构建良好的国内RISC-V创新环境

● 积极推动RISC-V标准化发展

2023年8月，中国电子工业标准化技术协会RISC-V工作委员会成立大会在京召开。RISC-V工委由中国电子工业标准化技术协会牵头，联合国内38家重点企业共同组建，是全国性、行业性、非盈利性的社会组织。工委围绕RISC-V产业发展开展标准研制、符合性评估、知识产权保护、人才培养、产业研究等方面工作，并与RISC-V国际基金会协同合作，引导产业实现优势互补、资源共享、协同推进，营造产业良好生态环境，带动产业链高质量发展。

● 国内企业成立RISC-V专利联盟

2023年8月28日，由中国RISC-V产业联盟、芯原微电子（上海）股份有限公司共同主办的第三届滴水湖中国RISC-V产业论坛在上海召开，论坛上9家RISC-V领域企业宣布成立RISC-V专利联盟。该专利联盟旨在构建RISC-V专利互不诉讼的生态系统，推动RISC-V技术创新和快速发展。首批联盟成员包括芯原股份、芯来融智、平头哥、赛昉科技、时擎智能、思尔芯、钜泉光电、芯思原以及上海恒锐知识产权共计9家单位。

（3）开源软件助力RISC-V生态发展

● 开放麒麟openKylin加入RISC-V国际基金会

2023年7月，开放麒麟宣布openKylin社区正式加入RISC-V国际基金会，成为其产业联盟成员，以构筑操作系统与RISC-V架构软硬件生态协同发展，助力RISC-V生态建设。openKylin是一个桌

面操作系统根社区，从源码级别自主构建了RISC-V开源桌面操作系统版本，目前能够支持HiFive Unmatched、VisionFive2、LicheePi和SG2042 EVB等多个硬件平台，支持80%以上市场主流RISC-V芯片，并与深度数智联合发布了全球首款搭载RISC-V笔记本电脑ROMA。同时，openKylin社区还基于RISC-V指令集和GPU优化了UKUI桌面环境，并适配了openKylin自研软件、Libreoffice、浏览器、影音播放和编程开发等常用软件。

● 开源欧拉openEuler加入RISC-V Landscape

2023年3月，开源欧拉openEuler宣布已加入RISC-V Landscape。openEuler以Community Organization Member的身份加入了RISC-V International，同时RISC-V正式成为openEuler的官方支持架构。RISC-V SIG领头完成了openEuler RISC-V 23.09的主线集成工作，该项目涵盖了主线代码回合、官方工程整合及正式的镜像的发布。自openEuler RISC-V 23.09版本起，RISC-V进入openEuler T1架构级别，这一进展标志着openEuler RISC-V发展的一个重要里程碑。目前开源欧拉已经推出多个基于openEuler开发的RISC-V架构的发行版。openEuler是由开放原子开源基金会孵化及运营的，面向数字基础设施的开源操作系统项目，支持服务器、云计算、边缘计算、嵌入式等应用场景。

3.4.4 中国与国际的RISC-V合作

(1) 合作推出RISC-V配套开源工具项目

2023年12月，中国芯原股份宣布与谷歌合作支持新推出的开源项目Open Se Cura。Open Se Cura项目是一个由设计工具和IP库组成的开源框架，配备了一个基于RISC-V ISA的开源、安全、低功耗的环境感知和传感系统，包括系统管理、机器学习和硬件信任根 (RoT) 功能。该项目旨在加速安全、可扩展、透明和高效的人工智能 (AI) 系统的发展。

(2) 合作推动RISC-V微控制器开发环境更加完善

2023年12月，中国兆易创新与德国SEGGER Microcontroller GmbH (以下简称“SEGGER”) 联合宣布，向所有使用GD32V系列RISC-V微控制器的用户提供免费商用的SEGGER Embedded Studio多平台集成开发环境 (IDE)。SEGGER Embedded Studio配备了强大的项目构建和管理系统、灵活的源代码编辑器，以及用于下载和安装的软件支持包，还集成了高度优化的运行时库emRun、浮点库emFloat以及智能编译链接器。目前SEGGER所有工具现已完全支持GD32V RISC-V MCU，包括

Embedded Studio集成开发环境(IDE)、市场领先的J-Link仿真器、Ozone debugger、实时操作系统embOS和通信、数据存储、压缩感知、物联网领域的软件库，以及Flasher编程器。

(3) 合作构建RISC-V车规级芯片的应用生态

2023年7月，瑞典嵌入式开发软件和服务公司IAR与中国RISC-V处理器IP及解决方案公司芯来科技共同宣布达成战略合作：IAR的Embedded Workbench for RISC-V车规级工具链将全面支持芯来NA系列车规级处理器内核，包括代码编辑、编译、调试等功能，以支持在芯来车规级内核中实现汽车功能安全(FuSa)。IAR Embedded Workbench for RISC-V是一个经过认证的集成开发环境，内含强大的IAR C/C++编译程序和调试器。

(4) 合作打造更丰富的RISC-V硬件开发环境

2023年7月，中国RISC-V设计企业平头哥与开源硬件开发社区BeagleBoard合作，推出首个兼容安卓的RISC-V单板计算机（SBC, Single Board Computer）BeagleV-Ahead。该单板机基于平头哥RISC-V内核芯片，开发者可以在机上运行安卓、Yocto、Ubuntu等多操作系统，能完成计算、存储及输入/输出等基础功能，并可直接嵌入智能产品中使用，可以协助开发者探索RISC-V在AI、物联网、机器人等领域的应用创新，降低了RISC-V开发门槛。

3.5 智能网联汽车行业研究报告

3.5.1 智能网联汽车行业发展概况

2023年1-12月，我国新能源汽车新车渗透率达到30.8%，纯电动乘用车和插电混动乘用车能耗指标基本达到2025年指标；2023年PA级智能网联汽车销量快速增长，达到800万辆，渗透率达到46.2%，年平均两位数增长，接近2025年50%的目标，进入技术驱动的快速发展期。目前智能驾驶技术的发展正在按照舱泊融合1.0（2023年）、舱泊融合2.0（2024年）、舱驾融合（2024~2025年）、中央电脑（2025年）的技术路线向前发展。随着电子电气架构从分布式向集中化，再到形成真正的车载中央计算机发展，自动驾驶汽车最终会形成多域融合的统一操作系统。统一的智能汽车操作系统能够提供更强大的硬件普适性，大幅降低整车企业在测试、维护上的成本，也能更好地推进车路协同的落地。

从软件技术层面，智能网联汽车的整体产业生态还存在很多挑战：

- 首先，自动驾驶操作系统内核、基础库、功能软件等核心底层技术方面，当前国内还处于技术架构探索和快速更迭阶段，相对来说差距还比较明显，行业未能在统一架构上推进开发，协同合作不足，存在力量分散和重复造轮子的现象；

- 第二，研发测试工具链目前国内的成熟产品不多，应用方在生产环境还不能做到大规模使用，难以形成市场化良性循环，也导致了研发企业无法获得需求方的原始反馈，技术与应用场景的代沟明显；

- 第三，车规级高性能芯片目前还没有发展起来，从计算芯片指令集、ISP/DDR/DSP等芯片IP，到芯片的设计、工艺还存在一定的差距，需要行业内集合一定的资源，逐步的摸索和解决。

汽车的软硬件生态需要从产业角度，整体性的构建。操作系统的底层生态与计算芯片、域控制器的发展分不开，研发和测试工具链的发展与系统集成、测试验证、产品化分不开，高速迭代的应用、好的用户体验与开放的需求和良性的商业环境、政策法规的发展分不开，要解决上面的三个问题，真正意义的进入软件定义汽车的时代，是一个行业性的集体课题。

3.5.2 汽车软件行业开源现状与典型案例

(1) 汽车软件行业开源趋势

软件定义汽车的时代，车用操作系统是智能网联汽车产业最基本、最重要的基石，直接决定了行业技术基础设施的能力水平。车用操作系统不仅管理着汽车的软硬件资源，芯片软硬件接口，同样管理大量数据资源，辐射带动相关传感器、软件算法、开发工具软件等领域发展；车用操作系统有效带动上游产业发展，是深化软件定义汽车、推动数字技术应用关键的基础设施，是汽车产业链和价值链的中枢。现在车用操作系统以商用闭源操作系统为主，采用黑盒交付方式存在开发局限，难以满足智能网联汽车应用软件的快速迭代需求。

预计未来三年将是智能汽车技术产业生态形成的关键期。然而车用操作系统作为一个复杂庞大的软件系统，从研发到应用是一个长期持续的过程，面临开发难度大、周期长、成本高等问题，需要持续投资，软硬件协同，挑战巨大。尤其是处于操作系统最底层的内核研发投入大、周期长，难以实现单独商业价值，单体企业难以承担。以开源共建方式推进车用操作系统的研发与生态构建已成业界共识。

(2) 汽车软件行业开源现状

随着汽车产业数字化转型深入发展，汽车软件扮演的角色愈发重要，加强基础研究，推动操作系统和基础软件国产化发展已成为行业共识。在软件定义汽车的趋势下，汽车行业发展的三个重要趋势之一就是从小源到开源。开源软件正以指数级增长，为汽车制造商提供了更为灵活、高效的开发环境，同时也推动了汽车技术的创新和进步。

操作系统内核属于共性基础技术，共用、共建、共同扶持是基本思路，一种方式是开源进行行业共建，一种方式是共同扶持头部企业进行开发实现共用。开源对于国家和产业来说是一个建立共同基础的重要途径，对于企业来说可以给开发带来更多的灵活度和自由度。

内核是车用操作系统的核心，提供系统最基础的功能，保障系统的性能和稳定。从内核开始向上发展，通过对基础服务、系统服务等持续建设，推动全面开放、社区开源，逐步发展出成熟的车用操作系统，更好的支撑行业生态建设。

当前，以百度Apollo自动驾驶开源项目、普华基础软件公司“小满”安全车控操作系统等为代表的开源项目正在快速发展壮大。与此同时，国内汽车软件生态正在不断快速发展中，如开源组织（中汽协、汽车工程学会、OpenSDV汽车开源联盟等）、汽车相关开源公司（如普华基础软件、国科础石、百度等）；典型开源项目（如Apollo、DDS、ROS等）、汽车软件开源组织形态（如极氪汽车、长安、小米等OSPO）等正逐渐形成更强大的开源生态。

作为一种汽车行业创新协作模式，开源不仅是开放源代码的软件技术开发，还包括更为广泛的开放技术领域及协同创新的理念与机制。目前汽车行业针对车控操作系统的开源生态建设已形成初步共识，后续有望分步实施、逐步演进：在操作系统产品发展期，开放是首先满足的条件，通过构建标准开放接口以支持更多行业伙伴产品的快速对接，实现更灵活的开发调试。在产品不断成熟期，可以有选择地开源部分代码以加速技术迭代，吸引更多生态伙伴参与开发和维护。产品完全成熟后，适时进行完全开源，形成稳定成熟的开源生态圈。

开源共建至少有四方面优势：

- 可有效降低车用软件开发成本，节省国外商业操作系统授权费用、工程服务费用；
- 降低车用软件开发入门门槛，保障软件开发可用资源，灵活使用行业可用软件资源、人力资源，提升软件开发效率及质量；
- 高效助力的技术发展与创新，使整车厂可聚焦技术创新及产品差异化竞争力打造。使用安心、可靠

的基础软件平台，整车厂可以更加聚焦体现技术创新及产品差异化的上层应用软件，避免行业内卷和垄断；

- 有效提升汽车软件产业能力，形成良好的产业发展生态及人才培养环境。使得懂汽车的人更容易懂软件，培养“既懂汽车又懂软件”的复合型人才。

目前行业已围绕开源进行积极探索，通过聚集行业力量，以更低的成本、更高效的模式共同打造车控操作系统、智能驾驶操作系统。中国汽车工程学会、中国计算机学会与国家智能网联汽车创新中心共建车用操作系统与泛在操作系统联合实验室，研发智能驾驶操作系统开源基线版本；OpenSDV汽车开源联盟已组织长安科技、国科础石、黑芝麻智能、Elektrobit、中科渝创等20余家单位，在该联盟操作系统工作组下发起及推进智能驾驶操作系统发行版项目，同时推进发展汽车实时操作系统、虚拟化、云原生等基础软件开源项目；中国汽车工业协会、开放原子开源基金会与中电科、普华基础软件等成立汽车软件开源工委，正在推进车控操作系统、中间件的共同研发。

开源共建不代表一哄而上，搞重复建设或者各自为战。开源项目的成败，取决于参与力量是否可以深度协同，打造完善的生态。主机厂更多会以提供工程化应用和验证的场景，基础软件开发领域的经验，科技企业结合工程应用需求，可参与内核核心代码、系统服务、基础服务、操作系统接口的开发工作，提升开源项目的实用性、适用性及通用性。未来，在共建开源操作系统的基础上，发动各方力量参与建立操作系统应用相关工具链、通用系统软件、功能软件、软件生态，社区的开发、建设工作。

考虑开源的不止中国企业。更早之前，博世和微软等企业也宣布成立“Eclipse软件定义汽车工作组”，为汽车行业提供开源的汽车软件，将复杂的汽车软件平台构建成一个巨大的生态。

2014年，Linux基金会发布了开源的AGL（Automotive Grade Linux，汽车级Linux）规范1.0版本，2018年丰田首次使用了基于AGL的车载信息娱乐系统，大众、戴姆勒、现代、马自达、本田、日产、上汽等主机厂也纷纷支持AGL。用于手机等移动终端的安卓（Android）操作系统是Linux最成功的发行版，由于其生态丰富、可实现用户习惯的无缝衔接，近两年广泛应用于国产自主品牌的信息娱乐系统，例如蔚来（NIOOS）、小鹏（XmartOS）等都是基于Android系统的定制化改造。

特斯拉Autopilot采用开源方案对Linux内核进行了实时性改造，分别实现了L2+级自动驾驶。此外，特斯拉CEO马斯克表示特斯拉可能会向其他电动汽车制造商开放部分汽车操作系统源代码，他认为这一做法能够对汽车行业提供帮助，等同于“安卓系统（Android）作为手机行业的通用标准”。产业界

有不少开源社区和商业公司在做Linux增强和完善的发行版,以满足智能汽车产品中的部分功能要求(比如AGL的Linux已用于Cluster & IVI功能,RedHat的CentOS Automotive/RHIVOS将用于ADAS功能)。相较于QNX,基于Linux改造而来的操作系统具有两大优势:一是Linux具备开源优势,可灵活地在各种CPU架构上运行,方便内核的切换;二是有丰富的软件库,可支持智驾和座舱功能的操作系统。借此,特斯拉避免从零开始搭建,能通过OTA进行快速修正与升级,不断提升用户体验、构筑竞争壁垒。

(3) 汽车领域典型开源项目

Apollo开放平台

Apollo开放平台是百度发布的向汽车行业及自动驾驶领域的合作伙伴提供的软件平台,于2017年正式对外开源。2023年12月19日,百度推出了全新升级版Apollo开放平台9.0。该版本在工程、算法和工具等方面实现了全面升级,通用层可赋能多种应用场景的规模化落地,整体操作更加灵活易上手,使用场景通用易拓展。极大提升开发效率的同时,可帮助更多开发者快速搭建属于自己的自动驾驶系统。其首次适配了ARM架构,并支持在orin下的编译和运行,在设备选型上为开发者提供了更为多样化的选择。

目前,百度Apollo自动驾驶开放平台已完成11个版本迭代,Apollo开放平台已汇聚了来自全球170多个国家和地区的16万余名开发者,拥有生态合作伙伴210余家,已成长为全球最活跃的自动驾驶开放平台。目前,社区累计发布内容超过1200篇,其中开发者贡献内容60余篇。此外,面向自动驾驶行业开发者的研发产品——Apollo自动驾驶开发套件(Apollo D-KIT),已为100多家高校、科研机构及行业伙伴提供研发支持。作为全球最大自动驾驶开放平台,Apollo拥有生态合作伙伴达210家,几乎囊括全球主流汽车制造商、一级零部件供应商、芯片公司、传感器公司、交通集成商、出行企业等,覆盖从硬件到软件的完整产业链。Apollo开放平台已在园区、矿区、港口、环卫、配送等低速自动驾驶领域得到广泛应用,其相关产品包括无人清扫车、无人快递车、无人售卖车、无人驾驶巴士、无人售货车、无人清扫车等。

Apollo开放平台于2023年12月27日正式升级了Apollo EDU高校计划,针对本科层次构建立体化校企合作人才培养解决方案,通过Apollo Studio开发者社区,百度构建了课程、实验、赛事一站式学习实践平台。截至2023年底,Apollo Studio开发者社区在线学习人次超过3.8万,累计覆盖全国开设自动驾驶相关专业类的理工院校超70%,985、211理工类院校覆盖率68%,成为国内自动驾驶在线培训规

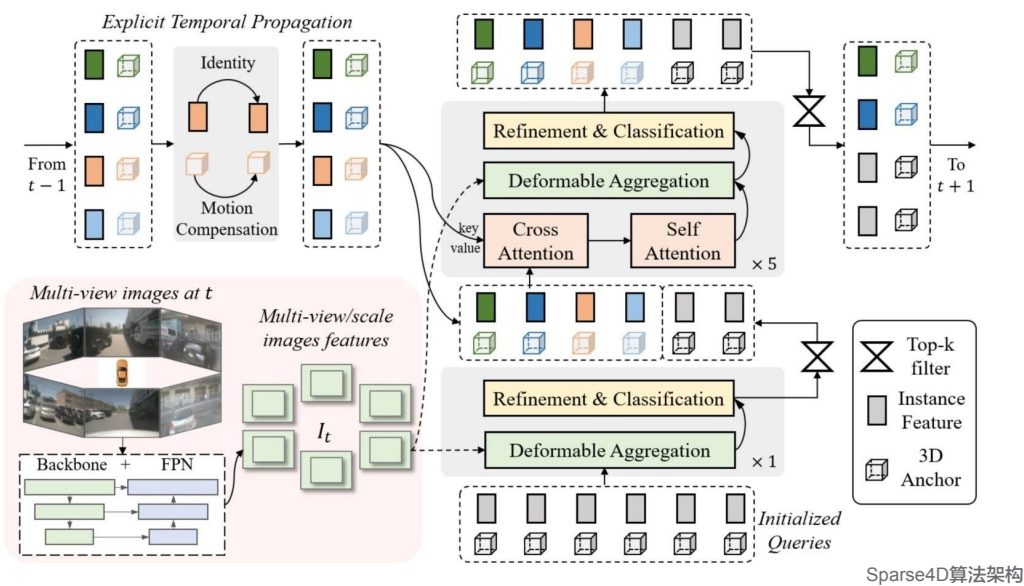
模最大和最有力影响力的社区。

中电科普华基础软件EasyAda “麓”，微内核开源项目

中国汽车工业协会正式发布中国车用操作系统开源计划中首个微内核开源项目，中电科普华基础软件向行业开源了“麓”（EasyAda）微内核源代码，可实现面向各类芯片平台和应用场景提供安全内核，具有内核微型化、服务模块化、遵循POSIX规范、可靠性高、实时性强、安全性高、可扩展性强等特点。采用木兰公共许可证（第2版），已在开放原子开源基金会代码托管平台AtomGit发布包括初始化代码、核心功能源代码、芯片适配代码等共计122个文件、14883行源代码。

地平线纯视觉自动驾驶Sparse4D算法

地平线将业内领先的纯视觉自动驾驶算法——Sparse4D系列算法开源，推动行业更多开发者共同参与端到端自动驾驶、稀疏感知等前沿技术方向的探索中。Sparse4D是迈向长时序稀疏化3D目标检测的系列算法，通过纯稀疏的融合感知框架，让感知算法效率更高、精度更高，让感知系统更简洁。相比稠密BEV算法，Sparse4D降低了计算复杂度，打破了算力对感知范围的限制，在感知效果和推理速度上，实现了对稠密BEV算法的超越。在nuScenes纯视觉3D检测和3D跟踪两个榜单上，Sparse4D均位列第一，成为SOTA，领先于包括SOLOFusion、BEVFormer v2和StreamPETR在内的一众最新方法。



地平线积极参与智能驾驶开源软件生态建设，此外还开源了基于矢量化场景表征的端到端自动驾驶算法VAD、端到端矢量地图在线构建方法MapTR等领先技术，持续推动行业加速发展。

国科础石础光操作系统

国科础石研发的应用于汽车智驾域、座舱域的础光Linux操作系统于2023年2月开源，在2024年初也获得了ASIL-B的相关功能安全认可，至今获得了诸多车企的良好反馈。础光Linux是国科础石在开源Linux基础上，针对汽车行业实时性优化改造的操作系统产品，该操作系统拥有更高的实时性及稳定性，有效增强功能安全性，并支持云原生。基于Linux的良好的软硬件生态，产品对NPU/GPU/Cuda等加速硬件可实现无缝衔接，减少厂商大量的适配和移植时间。

（4）开源社区及行业机构

中国汽车工程学会、中国计算机学会与国家智能网联汽车创新中心共建车用操作系统与泛在操作系统联合实验室，研发智能驾驶操作系统开源基线版本，推进产品持续迭代、验证。普华基础软件联合一汽、理想、长安等20余家单位结成开源共建伙伴，探索实施开源计划。

OpenSDV汽车开源联盟是于2022年9月在中国科学院软件研究所的牵头发起下，由汽车行业及相关技术领域的企事业单位、社团组织、高等院校、科研院所等参与组成的以推广开源开放促进汽车产业发展的非营利组织。

3.5.3 汽车软件行业开源的问题、风险和挑战

汽车行业与传统IT行业还是有较大的差异化，汽车非常注重安全合规，并在全球供应链协同下设计、生产和制造，开源对汽车网联安全的影响非常重要。从多次行业的交流来看，业界对于共性需求的底层操作系统、中间件、算法等技术底座，都倾向于“联合开发”，组团、联合等方式存在很多弊端，而开源软件模式则在传统IT行业有了很多的成功案例，采用开源开放的模式，共同解决通用的技术底座问题，是众望所归。

在汽车软件行业尤其是对于功能安全要求极高的车控、智能驾驶等环节的底层软件上应用开源模式，与传统基于开源的研发模式是有极大差异的。协同研发的社区基础设施、结合功能安全要求的代码审核机制、验证和仿真的平台，都需要面向汽车行业进行定制。这样才可以促进开源开放模式在汽车行业获得共识，便于开展工作。

此外，在垂直整合过层中利用开源模式，可加快软硬件协同效率，但也面临车企与供应商之间商业利益平衡分配的问题。

3.5.4 发展建议

长期来看，基于开源的智能网联汽车解决方案可能成为一种趋势，车企的自研路线将参考开源社区的技术。汽车行业软件基础设施层面会有大量开源方案出现（包括操作系统内核、中间件、框架、算法方案等），随着整车技术占比的提高，研发人员的多样化，车企研发团队也将与开发者社区建立深度连接，会出现开源、开放、协作式的汽车软件开发者社区，通过社区支持客户，沉淀知识，提升支持效率，也将成为一种产业生态建设模式。

从长远发展计，希望汽车行业有更有效社会化组织、能够开源结合开放，同时增强适当的基础设施建设，从而降低开发者参与门槛，加速产业和行业规模化发展，共建的开源开放项目，逐渐会形成一定的事实标准，与众多的顶层标准一起，螺旋式地促进技术的进步与产业发展。

3.6 开源供应链研究报告

3.6.1 开源供应链整体态势

开源供应链主要关注开源产业生态中不同角色间的依赖关系，以及根据上下游关系梳理出的核心节点。这些节点既包括供应链的主体如开发者、开源社区、开源基金会，开源企业，也包括开源代码、开源项目、开源产品，以及广泛存在的链接不同主体和客体的中间载体，如编译器、开发和测试工具、开发语言、包管理器、代码托管平台、软件仓库、测试用例等。当前全球开源供应链的发展已日趋成熟，形成了项目数量庞大、精细化、专业化运作的社区，各国也在不断完善开源供应链的治理体系。美欧等国家和地区将开源提升至战略层面，产业界成立基金会协调相关企业共同维护开源供应链安全，一系列新的开源供应链相关标准也在逐步探索实践，关于开源供应链的研究正在逐步推进。供应链安全风险日益严峻，风险类型可以划分为技术、法律和其他三大类。

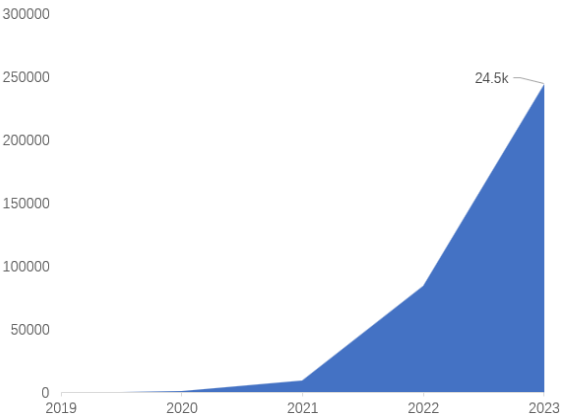
3.6.2 开源供应链面临技术风险

开源供应链技术风险主要表现在大量开源项目、开源组件存在安全漏洞。在开源软件供应链中，供应链上游的漏洞会随着组件依赖和代码复用被引入到下游的应用软件中，因此漏洞数量近年来逐年

递增，开发流程和最终使用者代码审计技能的不完善也会增大安全风险。直接和间接使用该开源代码或组件的用户基数越大，开源项目对攻击者的吸引力就越大，一旦出现问题影响也越大。

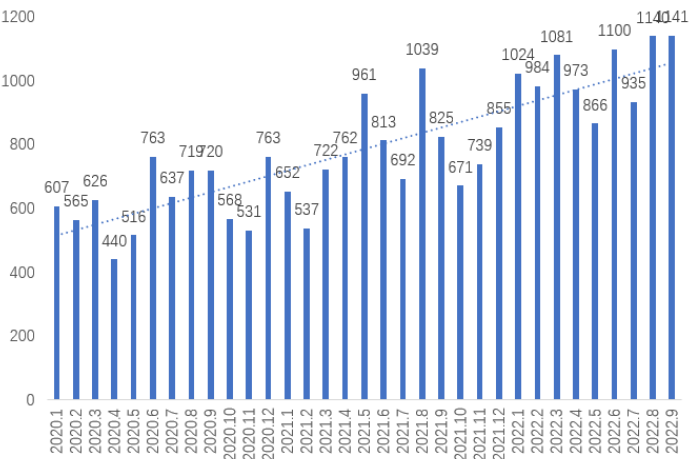
(1) 安全漏洞数量持续增长

在过去几年中，包含漏洞的开源软件数量持续高速增长。安全企业Sonatype《Annual State of the Software Supply Chain 2023》调查统计了390多万个开源库，其中有245032个存在漏洞，较前五年相比急剧增长。



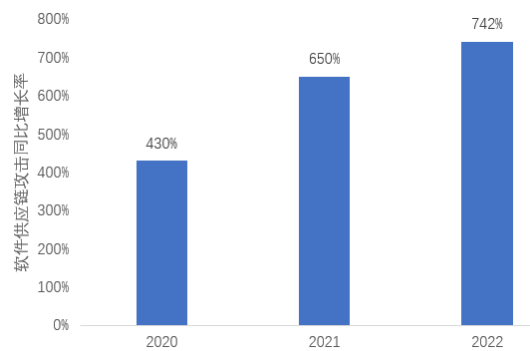
含有漏洞的开源库数量持续增长

已发布的开源软件漏洞数量也在不断增加，根据网安企业Mend最新一期的《Mend Open Source Risk Report》统计，近年来全球开源软件漏洞数量持续增加，在2022年，平均每个月新增开源漏洞数量已超1000，如下图所示。



近年来全球开源软件漏洞数量持续增加

漏洞增加的原因有很多，一部分原因是现在有更多的自动扫描和检测工具可以在开源软件中发现和修复这些安全漏洞，更重要的原因是全球利用上游开源生态系统（如JavaScript、Java、.NET和Python）的弱点而攻击的行为在持续增长。根据安全企业Sonatype的《Annual State of the Software Supply Chain 2023》调查统计，在2022年开源软件供应链攻击同比增长742%，较前一年的650%再次大幅增长。



2020年至2022年全球开源软件供应链攻击同比增加

开源软件供应链易的风险点在于攻击者发布恶意开源项目、植入恶意代码、使用易于合法包混淆的包名等。和其他软件安全风险相似的是，开源软件攻击者的主要手段是数据窃取、勒索病毒、获取主机权限、拒绝服务攻击以及其他获取经济利益的手段。根据新思（Synopsys）公司《2024年开源安全和风险分析报告》，Black Duck漏洞审计团队在2023年审计的1067个流行代码库中，96%的代码库包含开源代码，84%的包含至少一个漏洞，74%的包含高风险漏洞。分布数量前十的漏洞如下表所示。

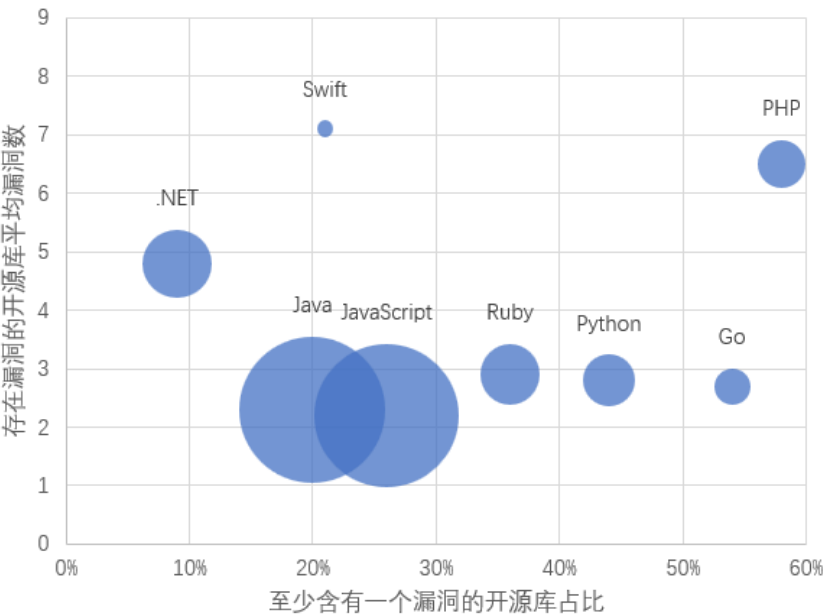
	漏洞编号	漏洞载体	漏洞类型
1	BDSA-2021-3651	jQuery	CWE-546可疑注释
2	CVE-2020-11022	jQuery	CWE-707对消息或数据结构的处理不恰当
3	CVE-2020-11023	jQuery	CWE-80Web页面中脚本相关HTML标签转义处理不恰当（基本跨站脚本）
4	CVE-2019-11358	jQuery	CWE-20输入验证不恰当
5	CVE-2015-9251	jQuery	CWE-79在Web页面生成时对输入的转义处理不恰当（跨站脚本）
6	BDSA-2014-0063	jQuery	CWE-79在Web页面生成时对输入的转义处理不恰当（跨站脚本）
7	BDSA-2015-0567	jQuery	CWE-937使用包含已知漏洞的组件
8	CVE-2019-8331	Bootstrap	CWE-97Web页面中服务端引用（SSI）转义处理不恰当
9	CVE-2020-7656	jQuery	CWE-80Web页面中脚本相关HTML标签转义处理不恰当（基本跨站脚本）
10	CVE-2022-31160	jQueryUI	CWE-79在Web页面生成时对输入的转义处理不恰当（跨站脚本）

分布数量前十的漏洞

(2) 安全漏洞层级传播

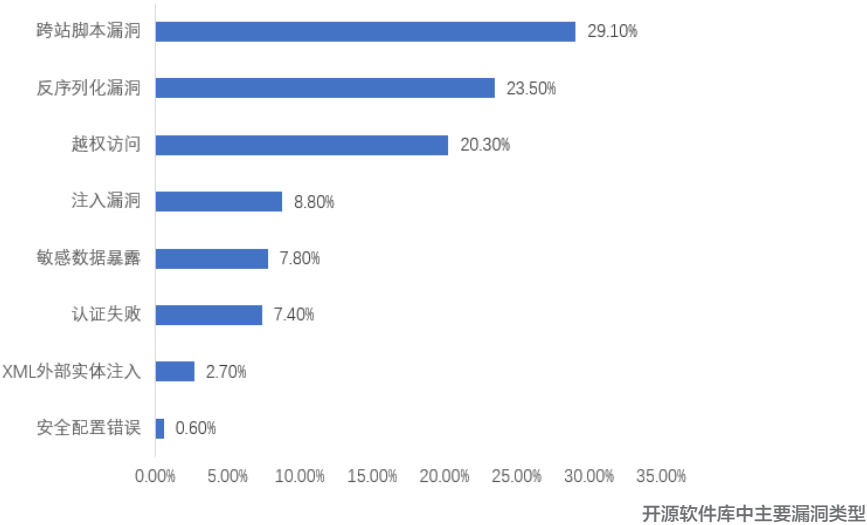
在开源组件复杂的依赖关系中，一旦上游项目发现有严重漏洞，就会直接或间接地影响到依赖它的下游开源软件，在错综复杂的层级依赖关系传播后，该漏洞隐匿在深层依赖的应用中不易被发现，为全球软件供应链带来无法估量且不可控的影响。

根据安全企业Veracode发布的《State of Software Security: Open Source Edition》，统计了8.5k个应用和35.1k个开源软件库的数据显示，软件应用中平均每个软件应用依赖283个开源库，有90%的应用依赖34个以上的开源库，有10%的应用依赖1400个以上的开源库。其中依赖情况最复杂的是JavaScript，JavaScript应用中平均每个应用依赖377个开源库，有90%的应用依赖66个以上的开源库。按编程语言统计，大部分开源库平均漏洞数量介于2-3个之间，包含缺陷库的百分比最低的是.NET，最高的是PHP语言。



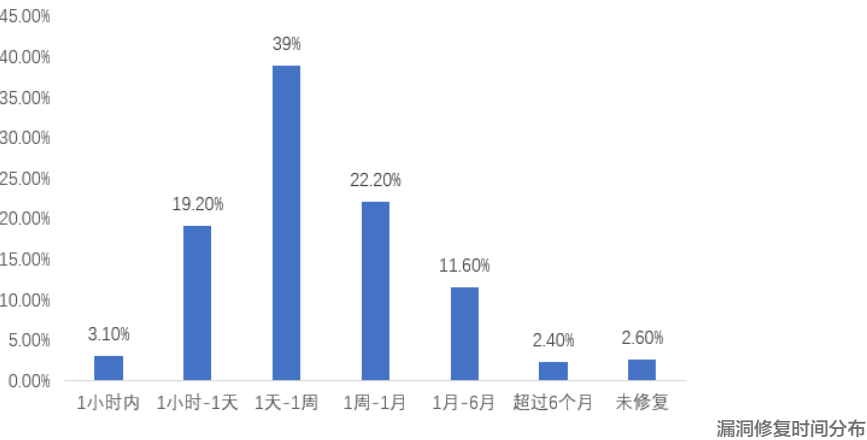
开源软件库中漏洞存在情况

统计这些库中所有已发现漏洞类型，比例最高的是跨站脚本漏洞，之后分别是反序列化漏洞、越权漏洞。前三种类型在所有已发现的漏洞中占比超过四分之三。



(3) 安全漏洞修复滞后

开源软件开发过程中使用过时的、存在漏洞的开源组件仍然是常态。据Linux基金会统计，超过80%的代码库包含四年未更新的组件。根据Sonatype的《Annual State of the Software Supply Chain 2023》调查统计，只有28%的组织机构可在1天之内追踪到最新发布的漏洞。39%的组织机构在1到7天内追踪到并着手修复漏洞，29%的组织机构需要1周以上的时间才能追踪到。这意味着将有很长一段时间，软件应用暴露潜在的恶意攻击下。



以2021年影响最大的Apache Log4j 2漏洞事件为例。据不完全统计，GitHub超过8600多个开源软件直接依赖Log4j 2组件，但通过这些开源软件继续追溯，最终超过20万个开源软件受到了影响。同时，在官方第一次发布修复版本的一周时间后，仍然有超过80%的间接关联开源软件没有被修复。

(4) 基础平台以及重要项目面临巨大的安全压力

2024年3月全球最大的开源软件平台GitHub遭受了大规模攻击，涉及克隆安全且干净的存储库，添加恶意的、模糊的代码后重新上传。据报道，此类攻击于2023年5月开始，呈指数级增长。随着这些攻击的继续，越来越多的用户可能被感染。

历史上，Github多次面临恶意攻击和信息泄露风险。例如，2023年GitHub平台上发生了大规模的敏感信息泄露事件，超过300万个公开代码库累计泄漏超过1280万个身份验证和敏感密钥。2018年，Github遭受峰值高达1.3Tbs/秒的严重的DDoS攻击，其他年份也多次遭受严重攻击。Github作为开源领域的重要基础平台，长期面临巨大压力。

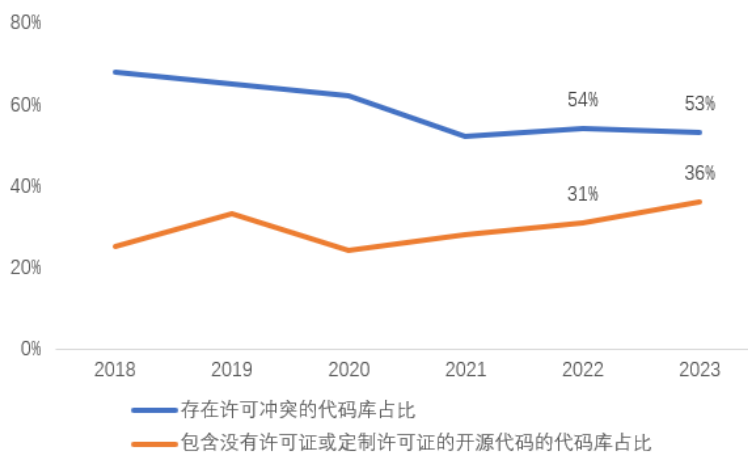
在重要项目方面，2024年爆出的开源软件xz后门事件引起业界震惊。一款在Red Hat和Debian等多个广泛使用的Linux版本中的压缩工具被发现植入了恶意代码，发现时该代码已存在三年，若xz被广泛引入各Linux发行版，Linux系统也将会轻易被入侵。目前，GitHub已经禁用XZ Utils代码库。

3.6.3 开源供应链面临法律风险

开源软件并不等于无限制的自由使用，必须在遵循相关开源许可证的基础上注重知识产权保护。常见的许可证有GPL、LGPL、AGPL、MPL、MIT、BSD和Apache等，各个许可证还包含不同版本。总体上可分为宽松式（Permissive）和责任传播式（Copyleft）两种。若不认真分析和甄别，很容易陷入许可证冲突，违反协议条款等法律风险。

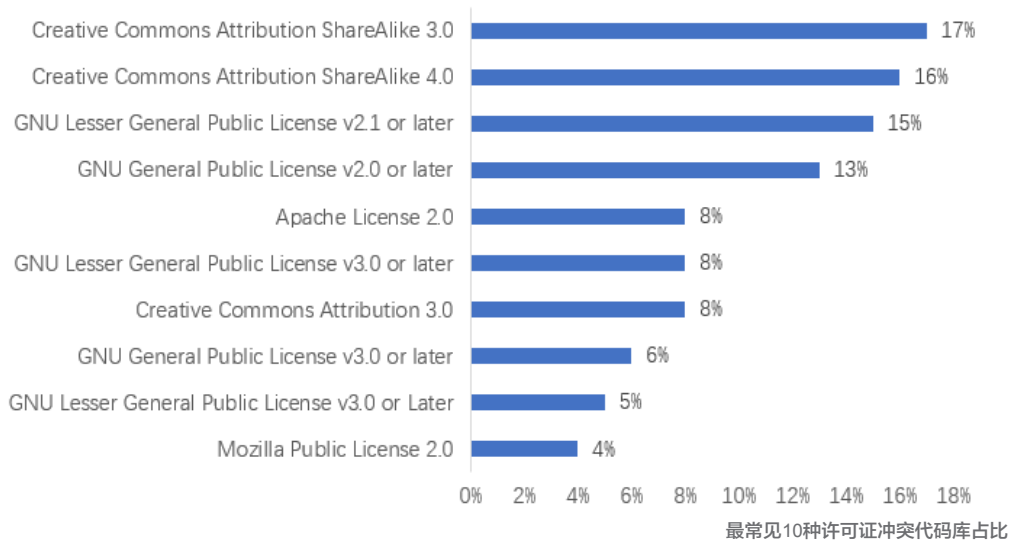
(1) 开源许可证冲突

信息系统中不同部分会存在多个开源软件许可证，但不同软件提供者对于软件发行的目的不同，因此开源软件许可证对企业或个人所给予的自由程度有所差别。而多个开源软件的条款之间可能会存在法律冲突，例如因署名义务引发的冲突，因发布义务引发的冲突，因许可证（GPL类）的传染性引发的冲突等，需要对具体代码库进行风险分析。根据新思（Synopsys）公司《2024年开源安全和风险分析报告》，Black Duck漏洞审计团队在2023年审计的1067个流行代码库中发现，53%包含存在许可证冲突的开源软件，有36%包含没有许可证或定制许可证的开源代码。



开源代码库的许可问题

其中，Creative Commons是引发许可证冲突的最主要的原因。仅Creative Commons ShareAlike 3.0(CC-SA 3.0)就引发了17%的许可证冲突。这可能是因为开发人员手动将代码片段或部分组件“复制粘贴”到代码库中，它们通常来自颇受欢迎的博客网站Stack Overflow，该网站根据Creative Commons ShareAlike许可证自动授权所有可公开访问的用户贡献。但这些许可证根本不是针对软件设计的，可能会引起法律方面的风险。



（2）特殊附加条款存在隐藏红线

即使有些企业在专利运营方面成果显著，被称作为“巨无霸”，但是在使用开源软件时也要注意一些“隐藏”红线。安卓系统是谷歌开发的智能手机和平板电脑操作系统，全球份额超过82%，虽然安卓系统被称为开源系统，但是开源系统也存在着“垄断”行为，最简单的例子是手机制造商必须接受预装的谷歌软件，否则将无法使用安卓系统。作为免费安装安卓系统的“对价”，如果一个设备制造商意图安装谷歌的任何服务，主要是安卓系统，谷歌都会在协议中要求他们安装其他的谷歌服务，在形式上非常类似于捆绑销售。这份“移动应用分发协议”（Mobile Application Distribution Agreement）（以下简称MADA协议）被其竞争对手在欧洲、韩国等地投诉滥用市场地位。分析MADA协议中的预装服务条款，一旦制造商决定安装一个类似于YouTube的谷歌服务，那么谷歌的搜索引擎就会被认为是搜索的默认引擎，谷歌地址服务也会默认安装并应用。由于“集群效应”的存在，制造商将有可能主动排除选择竞争对手的软件。

另外，有些软件被称为“开放源代码”软件但实质上同真正的开源软件有着本质不同，例如微软提出的“共享源代码计划”（Shared Source Initiative），该计划所提供的绝大部分源代码都允许合格者下载，并通过各种许可证进行规范和限制，而用户并不能直接对源代码有任何修改，只允许阅读源代码或提出修改意见。

（3）开源知识产权问题悬而未决

开源供应链还涉及到著作权，软件专利，商标和商业秘密等法律风险。首先，由于开源软件都有免担保条款，强调“在任何情况下，软件的原始开发者和版权所有者都不对任何损失负责，即使他们已经被告知这种损失的可能性”，但如果软件作者明知某代码存在危害性，同时未署名并标明修改内容，软件作者的这种行为并不会受到免担保条款的保护，依旧会被认定为侵犯了他人的知识产权。其次，开源软件的著作权仍未明确，世界各地的开发者都会参与进开源软件的改进，这就不会出现对开源软件进行修改且具有独创性，修改的结果使得开源软件能单独成为一个新的作品，该作品也受到著作权法独立的保护。再次，后续开发者将基于源代码产生的作品进行专利申请，导致使用该软件的用户可能会侵权专利权。最后，开源软件的开发人员之间并非以真实姓名联系，沟通的方式一般通过网络进行，因此开源项目组很难判断个人提交的代码是否侵犯他人的知识产权。

另外，随着近年来大模型持续火热，人工智能工具也使得开源代码知识产权变得日益复杂。随着AI驱动的编码推荐工具的使用，围绕生成代码的所有权、版权和许可问题也随之产生。例如，一项针对

GitHub、Microsoft和OpenAI的集体诉讼声称，GitHub Copilot违反了版权法和软件许可要求。具体包括Copilot推荐的代码使用了未经许可的内容，没有归属、版权声明，或者不遵守原始许可条款。该案例凸显出使用AI生成代码的法律复杂性。

3.6.4 开源供应链面临其他风险

阻碍开源供应链上下游高效整合的一大因素是信息不透明，展开来讲是开源组件的成分、依赖信息的缺失或紊乱。而全球范围内供应链安全标准体系仍未统一，专业性的供应链安全组织尚未建设成熟。维护开源供应链本身逐渐成为一项成本高昂的工作，这导致小团队的项目因能力有限而逐步停止维护，而大公司凭借资源优势易在细分领域形成生态垄断，这将对开源供应链的健康发展带来巨大风险。

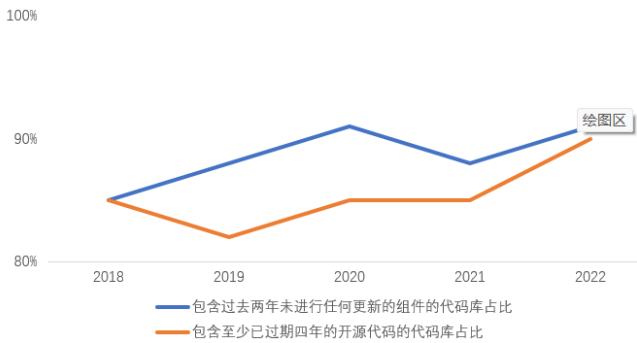
(1) 小项目停止维护

当一个开源代码库变得越来越受欢迎时，其代价就是其维护人员承受的压力就会越来越大，随之产生一系列运维相关问题。大量的项目处于失去维护的风险中，据奇安信代码安全实验室《2023中国软件供应链安全分析报告》数据显示，2022年全年，主流开源软件包管理工具生态系统中不活跃的开源软件项目数量为3967204个，占比高达72.1%。统计的八个典型的开源软件包管理工具生态系统Maven、NPM、Packagist、Pypi、Godoc、Nuget、Rubygems、Swift中NPM的不活跃项目数量最多，达1693287个，Rubygems的不活跃项目比例最高，占比达到90.5%。具体见下表。

序号	包管理工具	项目总数	不活跃项目数	不活跃项目比例
1	Maven	656090	451557	68.8%
2	NPM	2328687	1693287	72.7%
3	Packagist	382623	283408	74.1%
4	Pypi	438973	289468	65.9%
5	Godoc	511587	348457	68.1%
6	Nuget	576800	313449	54.3%
7	Rubygems	173691	157230	90.5%
8	Swift	90136	78548	87.1%

开源软件包管理工具生态不活跃项目数量

开源项目是由不确定数量的贡献者和维护者共同努力的结果。这种结构使开源成为一种协作项目，但由于开展维护活动的贡献者缺乏激励，小型团队常常难以支持开源代码的持续维护。新思《开源安全和风险分析报告2023》统计的1,481个代码库中，91%都包含过去两年内未进行任何更新的开源代码。这可能意味着相关方不再对该项目进行维护，尤其是在小型项目。



近年来开源代码库可维护性

2020年初，JavaScript的模块化标准库core-js的作者Denis Pushkarev因交通肇事被判处18个月监禁。Denis Pushkarev是core-js项目唯一的维护人员，该项目周下载量达2600万次，在npmjs.com上被19279个包依赖，重要程度可以认为是前端的基础设施。后续，GitHub社区通过将账户所有权转移暂时缓解这一问题。

2022年1月10日，个人软件开发者Marak Squires将其个人创建的位于项目仓库GitHub和开源组件包NPM上的开源库faker.js、colors.js的代码清空。由于成千上万的用户依赖这些库，本次恶意更新导致所有相关项目受到影响。使用遭到破坏的版本，会导致应用程序无限输出乱码。据报道，其清空仓库的代码是因为缺乏资金和被别人滥用开源项目，并声称不希望自己的努力成果为国际巨头企业免费使用。由于faker.js、colors.js代码库的历史版本仍存在于相关平台上，用户可及时降级为旧版本解决。后续，项目仓库GitHub还暂停了该作者对自己所有公共和私有项目的访问。

2022年1月11日，另一个开源项目Apache PLC4X的创建者Christofer Dutz也在开源托管平台GitHub中发文，称由于得不到任何形式的回报，他将停止对PLC4X的企业用户提供免费的社区咨询，若后续仍无企业资助项目则将停止项目维护和任何形式的支持。

(2) 大企业生态垄断

随着开源势不可挡地发展，全球科技巨头持续加码开源领域，通过收购开源平台逐渐强化其市场主导地位，并使得社区从早期的高度分散的技术架构转变为由几个强大的网络巨头所控制的架构。这种“大公司拥抱开源”的现象，一方面因大公司拥有在更高级别上开发和维护开源项目所需的资金，

推动产生了更多的开源重点项目，并有助于提高质量和安全性；另一方面，受商业利益等因素驱使，通过对开源社区项目的开发和商业化推广，对开发人员施加种种限制，易造成技术垄断，并最终阻碍技术创新。

例如一些国际巨头通过通用性和适用性的开源产品，在产业链中具有掌握极大话语权。这些开源产品自身具有一定通用性和适用性，便于后续开发者“不重复制造轮子”而在其产品的基础上进行进一步开发，但又通过人为制造诸如广告服务、有意破坏其他竞品的用户体验等方式排除竞争者获得垄断地位，最后获得高额利润。尽管这些产品本身都是开源的，但是中小企业和个体开发者在面对限定的服务接口、苛刻的商业条款等问题时，很少有能力再进行创新并反馈开源社区。

（3）云服务等新模式的影响

开源在云服务兴起和大规模应用过程中发挥了至关重要的作用，在云原生等新兴领域也将持续产生深刻影响。但另一方面，云服务也给开源产业带来了新的挑战和冲击。云提供商可以通过采用开源软件搭建相关服务，并使其商业化，但云提供商并不必须为这些开源技术买单或者支持未来开发。相关的争议引出了在开源社区已存在一段时间的问题。开源使用效果最好的地方是软件基础设施，而不是应用软件项目。如果云计算公司成为软件的基础设施提供商，它们的市场控制力可能让它们得以接管开源项目，并以低于销售开源服务的公司的价格销售这些软件服务。云服务商可以此获得大量利润，但对回馈社区基本没有贡献。

部分开源研发的企业对云服务商因开源获利却不回馈社区的行为难以忍受。例如数据库系统Redis就因此更改了协议。Redis不再使用BSD 3-Clause开源协议进行分发。从Redis 7.4版本开始，Redis将采用SSPLv1和 RSALv2双重许可证。Redis的源代码将通过Redis社区版免费提供给开发者、客户和合作伙伴。但托管Redis产品的云服务提供商将不再允许免费使用Redis的源代码。

（4）国际形势以及政治因素影响

近年来，美国不断修改《出口管制条例》（EAR）来试图限制和阻挠一些包括开源技术在内的先进技术出口到特定国家、特定实体和雇员，并先后采取了如账号禁止访问、代码删除等行为来实现“断供”。另外，开源软件供应链近年来也越来越多地面临意识形态、地缘政治、战争冲突等因素导致的影响。多个开源平台、开源项目要求开发者和用户在参与开源的时候进行表态，并根据政治立场划分阵营分裂对抗，社区中开源协作的基础被人为割裂了。虽然开源社区一向推崇“自由、平等、相互尊

重”的原则，开源精神被无数开发者奉为圭臬，然而“断供”“封锁”之类事件的不断上演，令全球开源开发者们开始担心，政治因素未来是否将逐步影响开源，“开源无国界”是否可能成为伪命题。

然而，从积极的一面看，并不是所有的开发者和开源社区都会轻易被政治倾向和意识形态所左右。在2022年3月，Eclipse社区中引发了一起线上争端，并引发大讨论，24小时内就被社区管理者和资深开发者平息。事情起因是由于一位来自欧洲的代码维护者，拒绝接纳来自中国开发者的一个补丁，补丁内容针对的是龙芯架构指令集。公开拒绝的原因既有技术原因，也有政治因素。不过，社区中绝大多数开发者选择保持中立，或要求不要在Eclipse社区中讨论政治话题，而一些中国开发者听闻之后，涌入该社区进行针对性抨击，争端有愈演愈烈的迹象。该社区主要负责人第一时间站出来平息事件，主要观点两个，这是一个开放的技术社区，任何人都可以有自己的政治主张，社区对此不予评价，但社区是一个进行技术交流合作创新的地方，社区不欢迎一切有关政治立场和倾向的话题在此讨论。同时，根据OSI对于开源定义的十个特征，Eclipse社区严格执行其中有关对任何人和群体，任何技术路线的非歧视原则，因此，社区不赞成拒收来自中国的补丁。通过这个案例，可以得出结论，开源本身是具有非政治倾向的一个群体，但其中不乏一些个体会有各种观点和主张，也会有偏见和过激的行为。作为中国开发者和开源爱好者，应该尊重开源“开放共享、平等协作”的创新精神，并坚定的支持开源社区远离政治因素的干扰，保护和建设这方净土。

3.6.5 开源供应链相关的前沿动态

正如上文所提到的，开源供应链安全以及相关的合规性是一个很难用简单语言表达的复杂话题。宏观而言，这项工作涉及跨多个学科完整而准确地执行多个步骤，并定期根据控制措施验证和审计这些结果。微观而言，在执行过程中，合规检测呈现一个“Left shift”的趋势，并与DevOps相结合。通过将合规性融入 DevOps 工作流程，开发团队可以节省时间，同时创建安全且低风险的代码。从而：

- 通过工具和平台强制执行管理，以保证合规过程的一致性。
- 自动化的收集合规信息并可视化展示。
- 对于合规内容，确保单一权威来源并最大限度的在组织内复用。

由于合规证据经常以文档形式保存，如何最大限度的减少文档重复工作，成为了一项挑战。值得注意的是，从数据格式出发，由于合规证明的部分文档的读者是人而非机器，在自动化的过程中，要同时保证数据对处理过程以及人机交互的友好性。

针对上述挑战，IBM开源的Compliance-trestle (Trestle) 项目进入CNCFSandbox。该项目以NIST 的 OSCAL作为工具和人员之间交换的标准数据格式，作为一种标准，允许团队在“治理”层和运营工具之间架起桥梁。

Trestle允许与其他合规性工具平台进行标准化联动。有助于基于 NIST OSCAL 标准自动创建和管理以机器可处理格式表示的各种合规性工件。OSCAL 标准提供了一个合规性框架和一组以机器可处理格式表示的相应关键合规性工件，使所有合规性文档都可以作为代码处理，从而以相同的方式进行处理和管理。

如OSCAL中的链接提供了一种引用任意资源的方法，允许引用本地或远程内容。链接可用于 OSCAL 模型中的许多关键字段。如引用 OSCAL 格式未表示的（外部）信息。这可能包括对（网络安全）法律法规的引用、对组织标准和指南的引用、对软件物料清单 (SBOM) 的引用等。

第四章 开源教育现状

开源教育旨在让数字时代的公民了解开源的理念与文化，增强数字化协作能力，共享开源开放的成果。开源教育涵盖开源文化教育、开源意识教育、开源技能教育三个方面。开源文化教育内容主要包括开源社区的形成机理和运转机制，以及开源活动得以持续生存和发展的机制、机理等，同时还包括开源历史、开源共识及开源的治理规则等。开源意识教育包括创新意识和开放透明的协作共享意识，创新意识表现在以开源技术迭代为基础，敏锐感知新兴技术的需求并进行快速创造，协作共享意识表现在适应全球分布式开发的趋势。开源技能教育体现为开发者传统的编程能力，开发者对开源技术和工具的使用能力，以及分布式环境下的协调、协作能力。

开源教育在当下发展迅速。其因技术进步和教育理念转变而兴起，特点是资源开放、社区协作与创新。优势在于降低成本、促进公平和提升质量。然而也面临资源质量不一、技术支持缺乏和教师培训需求等挑战。目前已在计算机科学等领域广泛应用，未来有望与新兴技术融合，更多机构和企业也将参与。总之，开源教育虽有不足，但前景广阔，为教育创新和公平提供了有力支持。

4.1 开源教育加速融入并贯穿教育全阶段

4.1.1 高、中、小学开源教育

在人才培养的高、中、小学阶段，开源文化与技术逐渐渗入到庞大的教学体系中。在“双减”政策颁布之后，中小学生也更有时间和精力去投入一些开源的项目，加强综合素质提升。虽然中小学信息化教育目前较少开设独立的开源文化与技能教学内容模块，但是由于开源具有低成本以及便于获取、学习和应用等优势，各地高、中、小学可以尝试以开源软件、开源硬件作为素质教育、STEAM教育、创客教育的工具，培养学生的科学精神和创造力。例如Arduino、Scratch等开源软硬件在儿童创客教育产品被广泛使用。与此同时，众多开源组织和开源在线教育平台组织了相关师资培训及学生创客活动，并提供相关系列课程和教具，支持学校进行开源启蒙教育。

4.1.2 高等院校开源教育

高校是开源人才培养不可或缺的重要基地，目前，国内很多高校已经在积极探索我国开源创新人才的培养路径，推动开源软件生态建设，提升软件人才与关键软件技术创新和供给能力。众多高等院

校加大开源基础设施投入力度，进行开源课程建设，组织开源相关讲座、竞赛以及多种线上、线下活动。例如，北京大学计算机学院周明辉开设的《OSS Development开源软件技术》课程，以学生实际参与的开源项目为线索，讲授开源软件开发涉及的理论知识、方法技术和工程实践，帮助学生了解开源软件开发、学习如何参与开源项目、训练开源思维。华东师范大学与同济大学共同创建X-lab开放实验室，开设开源软件通识基础课程，通过收集、整理、理解和拓展国际最新的前沿开源课程，讲解开源基础、开源经济学、商业开源软件、开源项目办公室治理导论的相关内容。北京信息科技大学设置开源软件开发技术课程，通过学习开源软件背景、开发方法等方面知识，使学生掌握开源软件开发要领和开源项目管理方式，了解开源软件成功的经验和失败的教训，学会利用开放源代码资源从事专业学习和科学研究，客观分析开源软件对社会、经济可持续发展的影响，鼓励学生积极参与开源软件开发并为之贡献自身力量。

4.1.3 职业教育与终身教育

开源职业教育与终身教育跟随行业需求飞速发展。近年来开源具有的越来越大的影响力，开源相关培训以及网络课程大量增加。越来越多的职业教育机构和培训机构开始引入开源软件和工具，为学生提供更加实用和创新的教育体验和机会。基于网络的大规模开放在线课程MOOC学习模式出现，大量优质开源课程资源免费共享，吸引来自全球学习者注册学习，其中典型的代表是edX、Coursera、Udacity等MOOC平台。同时，一些企业也开始注重开源软件和工具的应用和推广，为职业教育与终身教育提供更加实用和创新的教育资源和支持。GitHub、GitLink、CSDN、开源中国、头歌（educoder.net）等开源社区上的优质项目也为学习者提供了面向全球顶级开发者学习实践的通道。在国家相关政策的指导下，随着教育信息化的发展，开源教育正在融入从中小学教育到职业化教育的各个阶段。

4.2 开源教育平台和开源协作模式

4.2.1 基于开发工具和知识社区的开源教育平台

基于开发工具的开源教育平台支持大规模开源开发者分享开源技术、传播开源理念、推广开源项目。基于开发工具的平台典型代表包括Github.com、SourceForge.net、GitLink.org.cn、Gitee.com、GitCode.net、eudcoder.net、learnerhub.net等。

基于知识社区的开源教育平台提供了软件工程服务、讨论区、资源库等，形成了活跃的互动在线学习模式。平台典型代表包括Stackoverflow.com、CSDN.net、learnerhub.net等。其中，Stack Overflow由Jeff Atwood和Joel Spolsky这两位非常著名的Blogger在2008年创建，旨在提供一个与程序相关的IT技术问答网站。中国专业IT社区CSDN(Chinese Software Developer Network)创立于1999年，致力于为中国软件开发者提供知识传播、在线学习、职业发展等全生命周期服务。

4.2.2 基于开放实践和开源竞赛的开源教育平台

基于开放实践和开源竞赛的开源教育平台支持大规模开源开发者参与开源项目实训，该类平台以实践需要为导向，形成了产教一体化的在线学习模式。例如，早期国防科技大学毛新军教授团队基于社交化编程思想和群智共创理念，设计了群体化学习平台原型系统(learnerhub.net)，其主体包括群体化学习内容共创、群体化学习主题研讨、群体化学习问题解决、群体化教学过程外包四大部分，用以支撑开源活动和进行开源教育。头歌(educoder.net)是一个由国内知名高校、产业联盟和大型企业共同发起的新型开放在线实践教学平台。基于大规模开放在线实践(MOOP)范式，头歌将1000多种开源软件引入教研和教学活动，支持2800所高校建立的8000多门在线实践课程，涵盖计算机程序设计、软件工程、计算机系统、云计算、大数据、人工智能、电子技术、智能制造等所有专业方向的教学、实验、实训和科研活动，同时也支持老师和开发者按需协同开发在线实践项目和课程，为各类高校和社会各界提供高可扩展和高可定制的教学资源和实践环境。

软通教育面向信创技术等领域提供教育和资源服务，自主开发的信创智能交互实训教学平台，支撑高校在开源领域的实践类教学活动、过程管理、学生创新实践能力以及师资科研，为高校提供专业建设、实验室建设、教师发展及学生就业实训等一站式解决方案，辅助教师快速适应新型教学模式、优化配置教学资源，快速提升学生在开源技术等领域的实际操作能力及行业经验，最终实现高质量的人才输出。

CSDN在20余年的发展中，逐步建成了体系化、多视角的教育平台。由优秀博主设计的专栏和课程在IT行业各个领域都有覆盖。由CSDN和专家主导的技能树、编程比赛、专项训练为IT人士提供了循序渐进、自我驱动的培养体系。CSDN推出的训练平台和能力认证，致力训练开发者的硬核技能，整体设计融入“技术素养”“项目素养”“开源素养”和“职业素养”。广大在校生可以通过训练平台进行技能实训，从而提升工程化交付能力。平台设计以任务闯关为驱动，培养主动学习的习惯，现已培养了大量具备动手能力的IT研发工程师。能力鉴定方面，CSDN能力认证以真实项目为考题，采用真实场景(全程录屏、真人露脸)的考核方式，线上实时监考，所有代码与文档同步到GitCode。

效果评估方面，有专家评判结果，有据可查。通过CSDN的训练平台和能力认证，为IT大学生毕业上岗缺乏实践经验及软件工程师技术能力客观评估提供了有效的解决路径。2023年3月，CSDN发布InsCode，它是面向开发者的集编程，创作，分享于一体的全新共创社区。可提供基于云的IDE，支持多种环境和语言，无需下载安装，帮助在校生直接在浏览器中编码开发；完整的开发流程，也可为他们省去大量配置和概念，专注在创作本身；提供不同类型的项目内容模版，助力开发者想法到作品的快速落地；开放的作品分享社区，为编程学习提供边学习边实践的学习环境。

百度AI Studio是针对AI学习者的在线一体化学习与实训社区，集合了AI教程、深度学习样例工程、各领域的经典数据集、云端的超强运算及存储资源，以及比赛平台和社区，从而解决学习者在AI学习过程中的一系列难题，例如教程水平不一、教程和样例代码难以衔接、高质量的数据集不易获得，以及本地难以使用大体量数据集进行模型训练等。

竞赛类学习平台主要采用在线编程竞赛的形式促进学习者进行程序设计语言学习、算法设计、编程应用等。例如，Topcoder平台（topcoder.com）每个月都有2至3次在线编程比赛，参赛者可根据自己的爱好选用Java、C++、C#、VB或Python进行编程。Kaggle平台（kaggle.com）为学习者提供了举办机器学习竞赛、托管数据库、编写和分享代码的环境，近些年吸引了大量学习者的关注。头歌将全栈化智能在线实践技术和环境引入在线竞赛，有效支持在线实战化的全流程竞赛和活动。自2018年以来，中国软件开源创新大赛已成功举办五届。2023年第六届“中国软件开源创新大赛”在国家自然科学基金委信息科学部的指导下，由中国计算机学会（CCF）主办，西北工业大学、绿色计算产业联盟、CCF开源发展委员会联合承办。大赛面向国家“十四五”开源生态发展战略布局，聚焦“卡脖子”软件领域以及人工智能、大数据、芯片设计、物联网等前沿技术领域的开源软件，旨在为国内开源社区提供展示、交流、合作的平台，激发开源创新活力，培养开源实践人才，助力开源生态建设的高质量发展。

4.3 校企结合的开源教育协同模式

4.3.1 多源主体协同推动开源教育发展

政府引导：2020年6月5日，教育部办公厅、工业和信息化部办公厅关于印发《特色化示范性软件学院建设指南（试行）》的通知中提出，要以特色化软件人才培养为目标，以深化产教融合为途径，以改革创新为驱动，以特色发展为重点，深化软件人才培养模式改革，大力开展关键核心软件技术攻

关，促进软件生态体系建设，充分发挥软件人才培养对产业发展的支撑引领作用，推动我国软件产业实现由大到强的历史跨越。同时，开源教育也在对接教育部“计算机领域本科教育教学改革试点工作计划（101计划）”，提升软件人才与关键软件技术创新和供给能力。

高校、企业、社区、科研院所多方积极协同：随着开源教育从以开源软件教学为主的教育模式向理论与实践结合的综合教育模式的转变，一些企业与高校联动，形成校企联合建设实训基地。例如极狐（GitLab）免费向X-lab成员中的华东师范大学数据科学与工程学院师生提供极狐GitLab旗舰版使用许可。并且，X-lab实验室将极狐在开源文化、开源社区运营管理、DevOps等领域的通用知识纳入教材编写与人才培养计划。同时，X-lab利用极狐技术平台与开源社区资源，基于极狐GitLab平台研发“校园版”DevOps。

平台支持：2021年，中国计算机协会（CCF）开源发展委员会成立，作为中国计算机及相关领域具有广泛影响的学术团体，重点聚焦共同打造开源、开放、中立的产学研协同开源创新服务平台，推动探索学术共同体主导的开源发展新途径，助力我国开源生态建设的发展。在OSTech、华为、Linux基金会亚太区、Linux基金会开源软件学院、南方科技大学等企业社区及高校开源俱乐部的联合支持下，成立了高校开源社区联合体——开源高校联盟，它提供社区运营、技术、课程/认证、导师以及多方面资源和资金服务。

4.3.2 开源教育与产教融合

高等院校开源教育从学生、教师自发开源教育行为转变为院校行为，从高校教师的独立行为转变为与开源社区合作的共建行为。北京大学软件与微电子学院的荆琦提出“产教融合下的双轨制开源教学模式”，开发《开源软件开发基础及实践》课程，与RocketMQ、openEuler、PostgreSQL、OpenCloudOS、CloudWeGo、XIAOJU-SURVEY、DoKit&Hummer、OpenAnolis等开源社区合作，开源文化通识性教育和开源实践开发指导双轨同步进行。课程通过与开源社区合作，解决了开源技术方向多样性与教师个人研究方向单一性之间的矛盾，在一定程度上弥补了企业人才需求与高校教育之间的差距，被中国开源软件推进联盟推荐为示范课程。目前，国内很多企业及高校已经在积极探索我国开源创新人才的培养路径，推动开源软件生态建设，提升软件人才与关键软件技术的创新和供给能力。例如，头歌（educoder.net）平台构建了开源软件实践教学生态系统，致力于打造基于开源的覆盖通识课程、系统类课程、平台类课程的教学实验平台及相关课程资源和开源生态建设场景。达到提升学生实践能力、推广开源思想、培养与社会和岗位需求紧密匹配的目的。

4.3.3 开源创新活动丰富校企联动机制

丰富多样的开源活动为校企合作与开放创新提供了多层次的交流渠道，为产教研用搭建了高效的合作平台。

中日韩三国IT局长OSS会议。中日韩三国IT局长OSS会议暨东北亚开源软件推进论坛自2004年建立以来，在三国开源领域合作中发挥了重要作用。大会总结了近年来中日韩三国在开源技术与评估、人力资源开发、标准化与认证、开源人才共建等方面的成功经验，并联合产、学、研、用的全要素力量，共培开源沃土，助力开源产业创新发展。

开源中国开源世界高峰论坛。开源软件推进联盟每年举办开源中国开源世界高峰论坛，高峰论坛通过“现场论道+云端连线”的方式，邀请国内外开源领域的专家、高校学者、企业领袖、技术大咖等齐聚一堂，深度分享开源产业发展现状与趋势、技术与模式创新以及人才构建等方面的实践经验。

CCF青年精英大会。依托CCF在学术资源、产业资源和社会资源连接中的桥梁和纽带作用，中国计算机学会开源发展委员会探索产、教、研、用联动的开源创新模式，打造开源、开放、中立的产学研协同开源创新服务平台，举办CCF青年精英大会开源专题论坛，论坛邀请产业界、学术界以及科研领域的开源领军人物，共同分享和讨论如何通过产学研协同加速高校开源生态建设。

“CCF开源高校行”。它旨在高校加强开源文化宣传和开源人才培养，是CCF开源发展委员会年度重点校园推广活动之一。活动从2021年开始，依托CCF开源发展委员会执行委员，联合企业、开源社区和高校，围绕开源文化、技术、项目、应用等设立不同专题，走进高校，面向广大师生宣讲开源。“开源高校行”活动至今共组织了19场，活动参与的高校包括北京大学、清华大学、复旦大学、中南大学、国防科技大学、南方科技大学、上海交通大学、武汉大学、武汉理工大学、华南师范大学、南京大学、东北大学等；50多位业内专家老师进行了报告分享；形成了广泛影响和成功实践。

GitLink确实开源、木兰开源社区、开放原子开源基金会、飞桨开源社区、红山开源平台、Alluxio开源社区、OpenCloudOS社区、OpenHarmony开源社区等众多开源社区专家、学者与高校师生分享开源经验，共话开源。

开放原子“校源行”是旨在积极宣传推广开源文化，加快推进开源“进校园”的开源公益项目，项目通过推广开源课程、建设开源师资队伍、举办开源活动、资助设立开源社团、引导开源实践、构建培训认证体系等方式培育高校开源人才。

2023年4月，开放原子“校源行”在北京航空航天大学起航，目前已经在北京航空航天大学、清华大学、哈尔滨工业大学、哈尔滨工程大学、北京理工大学、电子科技大学、武汉大学、西北工业大学等67所高校成立开源社团或开展社团合作运营，并与近300所院校建立合作网络；累积在全国100多所高校开展开源主题活动300余场；构建了近1000人的开源通识、开源技术师资力量；上线了100余门开源课程，覆盖操作系统、数据库、大模型等核心技术领域；并筹备了OpenHarmony、openEuler、OpenTenBase等三个技术方向的认证考试，目前OFCA-OpenHarmony助理工程师认证考试已经通过16,000余人（截止2024年7月底）。

4.4 开源教育问题与挑战

4.4.1 国内开源人才短缺

开源教育需要大量的技术人才和教育专家来推动，缺少专业人才是我国开源教育发展道路上一个极大的绊脚石。虽然各大高校每年都向社会输送大量的软件人才，但是还是不能满足开源教育的需要。从芯片、操作系统、虚拟化、云计算、人工智能、代码托管等领域来看，国内仍然依赖大量海外开源项目，如何带领学生进入“卡脖子”技术开源社区，培养一批未来的核心贡献者团队；如何摆脱受制于人的困境实现科技自立自强，仍然任重道远。

4.4.2 开源教育资源不完善

国内开源教育仍然处于少数高校、教师先行示范阶段，并未大范围展开。在教学资料方面，现阶段缺少面向开源的教育资源，包括教材、课程设计等。如何形成统一的教学案例、教学素材，使全国范围内的大部分高校达到实际推行开源教育的地步，仍是一个亟待解决的问题。

4.4.3 高校对开源教育的接受程度有限

传统教育的观念较为保守，教师和学生都习惯于被动接受知识，而不是积极地参与到知识的创造和分享中，使得教师与学生对于开源教育的理念和模式理解不足。从学生角度来说，缺乏开放学习的方法，从教师角度来说，应用计算机的能力较低，对新技术的使用难以适应，从“不会用”、“不能用”到“不想用”，从而导致开源教育难以在教育体系中得到认可和推广，限制了其发展。

4.4.4 缺乏可持续发展的商业模式

国内开源商业模式虽然逐渐成熟，但在开源的商业策略、长远的战略上与国外的企业还有着较大差距，国内开源教育项目大多采用捐助、志愿者、增值服务、机构合作、会员等绝大多数非盈利的运营模式，缺乏可持续发展的商业模式，难以支撑项目的长期发展。因此，寻找开源教育运营规律，实现商业模式优化和创新，是开源教育持续发展的不竭动力。

4.4.5 法律法规的约束力度需加强

开源是一种有约束的开放共享，正是这种约束才使开放共享得以长久繁荣。从法律法规角度，对违背开源许可证的负面行为进行合法合规约束，是推动国内开源事业的关键。国外有对FaceBook、Google、微软等企业进行反垄断罚款的先例。国内在这方面刚刚起步，如何在社会范围内建立健全版权意识、如何合理合法的对开源负面行为进行追责，都是我们当前应该考虑的问题。

4.5 开源教育发展建议

4.5.1 加强培养开源创新人才

推动基于优秀国产开源成果的课程体系设计、师资队伍建设和培养计划制定，培养开源创新人才，支撑国产开源软件形成可持续发展生态。建议由教育部牵头，会同工信部等国家部委，以开源实践教学、新形态开源教育资源建设、“101计划”和特色化示范性软件学院建设为基础，推动基于开源的教育落实到相应方案和规划中。

4.5.2 加强开源文化和开源技能教育

我国高校亟需加强开源文化和开源技能教育，建立鼓励软件成果开源的评价机制和价值导向，推动高校产出更多原创性开源成果。一是在软件工程等相关课程内容设计上，主动融入开源文化、开源规律和开源模式的教学内容，培养学生的开源意识和开源技能。二是在高校学科建设评价机制方面，将成果开源情况纳入考核指标，鼓励高校师生积极参与开源社区贡献，加强与开源企业的合作，促进开源教育的应用和推广，鼓励高校形成原创性开源成果。

4.5.3 打造创新性产教融合平台

开源教育应拓展到人才培养成长的各个阶段，需要相应平台和政策环境的支持，建立从高校开源学习、产业开源实践到开源创新创业的闭环，服务我国软件产业发展。一是以各类教学服务平台为基

第五章 开源企业及机构发展现状

5.1 中国开源企业发展现状

中国开源企业开源贡献情况按照民营科技类企业、国有企业选取部分典型，科技类企业包括华为、阿里巴巴、蚂蚁、滴滴、京东、九州未来、平凯星辰、思斐、涛思数据、统信软件、微众银行、小米、易捷行云、睿赛德、中兴通讯，国有企业包括中国联通、中国移动研究院，其他包括开源基金会与COPU、北京开源芯片研究院、北京智源人工智能研究院等其他组织。

5.1.1 科技类企业

(1) 华为

华为积极拥抱开源软件开发，目前是Apache软件基金会、Linux基金会、Eclipse基金会、开放原子开源基金会、OIF基金会、CNCNF基金会等数十个国际开源基金会的顶级成员或创始成员，并担任数十个董事席位以及数百个TSC、PMC、PTL、Maintainer、Core Committer等核心席位，在全球开源社区中积极贡献，代码综合贡献位居全球前十，亚太首位。近年来，华为在操作系统、数据库、AI计算框架以及云原生等基础软件领域积极贡献。由华为贡献和发起20+开源项目，已吸引成千上万的开发者，获得广泛认可。



华为发起的或作为主体维护的开源项目

OpenAtom OpenHarmony开源项目（简称OpenHarmony）是由开放原子开源基金会孵化及运营的开源项目，由开放原子开源基金会的OpenHarmony项目群工作委员会负责运作。OpenHarmony遵循Apache 2.0协议开源，截至2024年3月26日，OpenHarmony位居Gitee活跃度指数第一名，已有7200+社区贡献者、160万+社区用户，开发者下载次数1亿+；拥有43个软件发行版、201款开发板/模组、279款商用设备，是码云平台上当前代码和社区最活跃的开源项目；覆盖金融、交通、教育、政府、安平、能源、制造、卫生、广电、电信、航天等领域，OpenHarmony社区已成为“下一代智能终端操作系统根社区”，携手共筑万物互联的底座，使能千行百业的数字化转型。

openEuler是由开放原子开源基金会孵化和运营的开源项目。openEuler是面向数字基础设施的操作系统，支持服务器、云计算、边缘计算、嵌入式等应用场景，支持多样性计算，致力于提供安全、稳定、易用的操作系统。通过为应用提供确定性保障能力，支持OT领域应用及OT与ICT的融合。自欧拉2019年12月宣布开源，截止2024年3月，欧拉开源社区已吸引上万名开发者，105个SIG组，超过1500家企业加入社区，汇聚处理器、整机、基础软件、应用软件、行业客户等全产业链伙伴。国内主流的系统厂商均推出基于欧拉的商业发行版，应用于政府、运营商、金融、能源、电力、交通等行业核心系统。截止2024年3月，欧拉商用累计已经突破610万套。

昇思MindSpore是新一代覆盖端边云全场景的开源AI框架，旨在开创全新的AI编程范式，降低开发者门槛，为开发者打造开发友好、运行高效、部署灵活的AI框架，推动人工智能生态繁荣发展。同时，昇思在致力于技术创新的同时，还着力打造学习型社区环境，希望凝聚开发者力量共建社区，与开发者共同学习和成长。截止2024年2月，MindSpore社区下载量687万+（2020年3月28日开源），社区核心贡献者25000+，服务企业数量5500+，与360+科研院所展开合作（教学290+）；开源模型500+（已发布10大模型，10+模型联合开发中），顶会论文1200+，全球排名第二，国内第一。

KubeEdge是业界首个云原生边缘计算框架，云原生计算基金会（CNCF）唯一孵化级边缘计算开源项目。KubeEdge完整的打通了边缘计算中云、边、设备协同的场景，为用户提供一体化的云边端协同解决方案。2018年11月开源以来，KubeEdge社区已累计发布70+特性版本，收获7k+Star，全球有1500多名开发者参与代码贡献，超过100家企业与科研机构参与项目合作。项目目前已被应用于广泛应用智能交通、智慧园区、工业制造、金融、航天、物流、能源、智能CDN等行业，完成业界最大规模云原生边云协同高速公路项目（统一管理10万边缘节点/50万边缘应用）、业界首个云原生星地协同卫星、业界首个云原生车云协同汽车、业界首个云原生油田等项目。

Karmada (Kubernetes Armada)是CNCF首个多云多集群容器编排项目。结合了华为云多云容器平台MCPI以及Kubernetes Federation核心实践，让用户像使用单个集群一样轻松管理跨云多集群，让基于Karmada的多云方案无缝融入云原生技术生态。自2021年开源以来，截至目前社区已获得4K Star，吸引了来自华为、DaoCloud、浙江大学、滴滴、腾讯、网易、Intel、IBM、Red Hat、Comcast等60多家公司的贡献者，广泛分布于22个国家和地区。Karmada现已在华为云、工商银行、兴业数金、中国移动、中国联通、携程、vivo、道客、飓风引擎、VIPKID、有赞、哗啦啦、之江实验室等众多企业单位生产应用，为企业从单集群到多云架构的平滑演进方案。

Kurator是业界首个开源的分布式云原生平台。通过集成Karmada、KubeEdge、Volcano、Kubernetes、Istio、PRometheus等业界主流开源技术栈，Kurator为用户提供跨云跨域多集群统一编排，统一调度，统一流量治理，边云协同，统一监控运维等核心能力。在信通院发布的2023下半年可信开源领域最新评估成果中，Kurator因在许可证合规性、软件安全性、软件活跃度、技术成熟度、服务支持力和软件兼容性多方面的卓越表现通过成熟度能力检验，获得中国信通院可信开源项目认证。

Volcano是业界首个云原生批量计算引擎，也是CNCF首个和唯一的批量计算项目，于2022年4月正式孵化。Volcano主要用于AI、大数据、基因、渲染等诸多高性能计算场景，支持几乎业界所有主流计算框架，加速整个云原生落地的进程。截至2024年3月，社区已吸引2.7万开发者，获得3.7k Star和800+Fork，海内外生产落地用户广泛分布于互联网、先进制造、金融、生命科学、科研、自动驾驶、医药等行业。Volcano作业管理能力被写入由中国信息通信研究院牵头制定的《高性能计算(HPC)云平台标准》，成为行业标准。

Kmesh是业界首个内核级云原生流量管理引擎，通过基础软件创新，帮助用户构建云原生场景下的高性能通信基础设施。基于eBPF+可编程内核技术，Kmesh将流量管理下沉到操作系统，消除了数据路径上对代理层的需求，并实现了内核级的sidecarless网格数据平面。通过将L4、L7流量治理能力卸载到内核，使得服务转发性能分别提升50%、60%，底噪开销降低70%，为应用程序提供了极致转发性能体验。

Kmesh在早期版本开发过程中，得到了openEuler社区的孵化与支持，后续作为独立发展的开源项目，将持续与openEuler紧密协作，为用户提供极致性能的流量治理技术方案截至2024年3月，Kmesh已经发布v0.2版本，用户可以在Kubernetes环境中使用helm一键部署，无缝连接到Istiod。Kmesh关键特性包括高性能、低开销、安全隔离、无缝兼容等。

Kuasar是一款云原生多沙箱容器运行时，2023年12月正式加入CNCF。目前，Kuasar已支持基于轻量级虚拟化技术的安全容器沙箱（Cloud Hypervisor、Qemu、StratoVirt），基于新兴的WebAssembly沙箱（WasmEdge、Wasmtime），基于进程级虚拟化的App Kernel沙箱（Quark）以及基于内核的原生普通容器沙箱（runC）。在北向引擎方面，Kuasar已与Containerd联合构建最新的沙箱接口标准，并共同推动该标准在Containerd v2.0版本的完整实现。

openGemini一款面向IoT和Devops场景垂直优化的云原生分布式时序数据库。主要聚焦于海量时序数据的存储和分析，通过技术创新，降低海量时序数据存储成本，简化系统架构，提升时序数据存储和分析效率。自2022年开源以来，已累计在60余家企业测试和生产落地使用，广泛应用于航空航天、医药、交通、车联网、电力、工业互联网等领域。在信通院发布的2023下半年可信开源领域最新评估成果中，openGemini获得中国信通院可信开源项目和可信数据库两项认证。

华为及所属人员获得的开源社区/项目荣誉（2023年度）

奖项	获奖项目/组织
2023 优秀开源人才培养企业（LFOSSA）	华为
2023 CCF 科技进步特等奖	OpenHarmony
2023 年度生态开源项目（开放原子开源基金会）	OpenHarmony
2023 最具价值开发者社区（CSDN）	OpenHarmony
2023 年度优秀开源技术团队（开源中国）	OpenHarmony
2023 开源创新榜 2023 年度优秀开源项目（中国科协科学技术传播中心&中国计算机学会&中国通信学会&中国科学院软件研究所）	OpenHarmony
2023 年东北亚优秀开源项目	MindSpore
第六届数字中国建设成果展览会“十大硬核科技”	MindSpore
2023 开源创新榜 2023 年度优秀开源社区（中国科协科学技术传播中心&中国计算机学会&中国通信学会&中国科学院软件研究所）	MindSpore
2023 生态开源项目（开放原子开源基金会）	MindSpore
2023 中国开源创新大赛一等奖（中国计算机学会）	openGauss
2023 软件行业活跃开源社区	openGauss
2023 最具影响力数据库	openGauss
2023 年度技术卓越奖	openGauss
中国信通院可信开源项目认证	Kurator 、 Sermant 、 openGemini 、 DevUI 、 OpenTiny Vue

中国信通院可信开源社区	KubeEdge 、 Karmada 、 Volcano
2023 开源创新榜 2023 年度优秀开源社区 (中国科协科学技术传播中心&中国计算机学会&中国通信学会&中国科学院软件研究所)	Volcano
“2023 快速成长开源项目”奖项 (开放原子开源基金会)	openGemini
“2023 快速成长开源项目”奖项 (开放原子开源基金会)	Kuasar
2023 开源之夏最具潜力奖课题 (中国科学院软件研究所)	KubeEdge
开放原子云社区副主席 (开放原子基金会)	王泽锋
CNCF 中国本土唯一 TOC 委员	王泽锋
CNCF Ambassador 2 席	王泽锋、任洪彩
2023 年度技术引领人物 (稀土掘金)	王泽锋
2023 开源创新榜优秀开源人物 (中国科协科学技术传播中心&中国计算机学会&中国通信学会&中国科学院软件研究所)	王泽锋
Apache Member、Apache Spark PMC、PyTorch TAC	姜逸坤
PyTorch2023 优秀贡献者提名	李佳伟
PyTorch Board	李永乐
2023 开放原子的开源活力贡献者技术、生态贡献奖	涂旭辉、曾令卡
2023 开放原子的开源活力贡献者生态贡献奖	齐宝玮、查浩浩

华为参与贡献最多的开源项目排名

华为积极拥抱开源软件开发，持续贡献主流开源项目。截止2024年3月，贡献开源项目排名如下：

- Kubernetes代码贡献量全球排名前十；Kubernetes Maintainer席位数亚洲排名第一
- Istio社区代码贡献量亚洲排名第一；Istio Maintainer席位数亚洲第一
- Linux Kernel 6.1厂家代码贡献排名第一，Linux Kernel总代码贡献量全球排名Top7
- Onap代码贡献量全球排名Top3
- Anuket代码贡献量全球排名Top3
- Akraino代码贡献量全球排名Top4
- DPDK代码贡献量全球排名Top6
- PyTorch代码贡献量全球Top7，基金会高级会员

- Spark近三年代码贡献量全球Top6
- OpenHPC代码贡献量全球Top3

(2) 阿里巴巴

阿里巴巴的开源项目

阿里巴巴持续操作系统、云原生、数据库、大数据&AI等多个领域自主开源项目，拥有超过3万个贡献者，位列中国企业开源活跃度榜首。

Anolis OS：一款开源、安全、稳定的服务器操作系统，全面兼容CentOS生态。

Nacos：一个更易于构建云原生应用的动态服务发现、配置管理和服务管理平台。

PolarDB-X：兼容MySQL的云原生分布式数据库领导者。

PolarDB-PG：采用存算分离架构、兼容PostgreSQL的云原生数据库。

DataX：它是阿里云DataWorks数据集成的开源版本。

Arthas：Alibaba Java Diagnostic Tool Arthas/Alibaba Java诊断利器Arthas。

SREWorks：Cloud Native DataOps & AIOps Platform，这是一个云原生数智运维平台。

阿里巴巴的技术贡献

开源作为阿里巴巴技术战略之一，阿里持续向头部开源基金会捐赠开源项目，作为开源生态中的领导者、实践者持续发挥力量。阿里巴巴持续向Linux基金会、Apache Software Foundation、CNCF基金会捐赠开源项目，基于阿里巴巴集团的业务场景，孵化出Apache Dubbo、Apache RocketMQ等顶级开源项目。

(3) 百度

百度自2017年起成为Linux Foundation、Apache Software Foundation、CNCF的黄金会员，也是开放原子开源基金会的创始白金捐赠人，共有10个开源项目捐赠给国内外开源基金会孵化，在开源社出品的《2023年中国开源年度报告》中位列中国企业排名前三。

由百度共发起的开源项目约1000余个，其中较活跃的开源项目有近百个。较知名的有

Apollo、PaddlePaddle、XuperChain、PaddleOCR、PaddleNLP、Amis、San、BaikalDB、Dperf、OpenRasp、BifroMQ、Apache ECharts、Apache Doris、Apache bRPC、Apache HugeGraph(Incubating)、Apache Teaclave(Incubating)等。其中三个项目Apache ECharts、Apache Doris、Apache bRPC已毕业成为Apache软件基金会的顶级项目。

百度及所属人员获得的开源社区/项目荣誉（2023年度）

- 飞桨开源社区入选中国科学技术协会发布的2022年“科创中国”系列榜单-开源创新开源社区。
- 飞桨荣获“2023中国计算机学会确实开源编程夏令营（GLCC）优秀社区”。
- 百度飞桨与开放原子开源基金会联合办赛，荣获“优秀共建方”称号
- Apollo社区荣获中国计算机学会工程师文化日贡献单位。
- 百度超级链XuperChain入选福布斯全球区块链50强。
- 百度开源办公室产品运营经理马红伟荣获思否“2023中国开源先锋33人”。

（4）京东

京东2018年以白金会员加入CNCf，2021年成为开放原子开源基金会白金捐赠人，是开源安全委员会委员，是开放原子开源基金会云社区工作委员会单位，也是OpenHarmony项目群工作委员会成员。

京东发起的或作为主体维护的开源项目

ChubaoFS，2020年发起，后改名为CubeFS开源项目，并成为CNCf沙箱项目。

Vessel kernel，2023年与openEuler社区积极共建在离线混部项目，向openEuler社区贡献云舰内核混部项目Vessel kernel。

Vearch是2019年发起的分布式向量数据库。百亿规模向量存储和查询，毫秒级耗时，专注于向量数据存储与查询的高性能、高可用和低成本。广泛应用于人脸识别，图片、音视频搜索、推荐系统及智能安防等场景中。在2023年5月正式捐赠给开放原子开源基金会，成为开放原子开源基金会正式孵化项目。

Taro开源项目是一款跨端的前端开发框架，在国内外的前端开发领域中都具有一定的市场份额，在

GitHub上Star数量达到34,600+, 累计服务超200,000开发者。

NutUI是京东风格的移动端组件库, 提供基础组件 (80+) /大促活动组件/抽奖小游戏组件等组件, 支持Vue3、React、RN等技术栈, 完成对H5、Taro多端、RN原生等适配, 帮助开发人员快速开发用户界面, 提升开发效率, 改善开发体验。

京东及所属人员获得的开源社区/项目荣誉 (2023年度)

- 基于openEulerOS研发京东云舰OS, 荣获 “2023年度openEuler领先商业实践奖”
- 积极参与openEuler社区开发工作, 深度参与openEuler kernel荣获2023年 “openEuler优秀项目奖” ;
- 荣获 “2023年度openEuler贡献之星” ;
- 2023开放原子开发者大会, NutUI作为开源项目同时荣获 “2023快速成长开源项目” 和 “2023活力开源贡献者-技术贡献” 的奖项;
- Vearch向量数据库项目核心开发者获得2023开放原子开发者大会年度开源贡献之星。

京东参与贡献最多的开源项目排名

- 2020年12月, 京东加入OpenHarmony项目工作委员会, 开始对OpenHarmony项目进行开源社区治理;
- 京东零售、京东科技、京东物流分别在OpenHarmony社区成立和主导了Taro的OpenHarmony适配子项目CrossPlatformUI SIG、OLA联盟开源组OH-OLA SIG、自动化机器人子项目Robotics-SIG;
- 参与OpenEuler Kernel社区,全年累计评审与合入PR数除华为外第一。

京东发起或捐献的最具影响力的开源项目年度进展

至2023年末, NutUI开源类产品累计Star 6900+, NPM下载量全年站外激增15.8w, 是去年增长的2.3倍; 站内增长2500+, 而其中React增加7000+, 2023年的下载量为22年的6倍。

Vearch向量数据库项目2023年度发展迅速, 迭代发布了9个正式版本, 实现云原生托管, 在大模型生态领域实现对主流集成框架的支持, 也推出了自动恢复和数据迁移等重大功能特性, 在托管平台上

PR、issue、Fork和Star分别增加到208、571、314、1.9K。外部开发者增加20人，开源用户增长上万，促进了国产AI原生向量数据的发展和应用。

Taro项目本年度发展迅速，推出正式版本30+个，在托管平台上PR新增000+个、Issues新增1500+个、Fork新增280+个、Star新增3000+。

(5) 滴滴

滴滴发起的或作为主体维护的开源项目包括Mpx、Dokit等。

Mpx是一款致力于提高小程序开发体验和效率的增强型小程序框架，目前在滴滴有着广泛的应用，支撑了滴滴的整个小程序生态。

Dokit是一款面向移动端的研发效率工具集合，主要为日常开发和测试提供各种便捷的辅助功能，提升开发和测试效率。中国科协科学技术传播中心等共同发布的“2023开源创新榜单”。

(6) 九州未来

九州未来十余年开源历程中，始终秉承拥抱开源、回馈社区的初心，始终致力于技术钻研和社区贡献，长久保持社区贡献率稳步增长。九州未来拥有OIF开放基础架构基金会董事会席位/技术委员会席位、国际第一个MEC边缘计算开源项目EdgeGallery董事会席位/技术委员会席位、边缘计算开源项目StarlingX全球技术委员会席位、全球边缘计算AKRAINO创始单位之一、中国开源云联盟COSCL理事单位、工信部开源计算标准制定单位等，以自身优势和积累持续为开发者提供开源价值、助推开源技术创新落地。

九州未来发起的或作为主体维护的开源项目

近年来，九州未来发起主导了建木、Skyline、OpenV2X、KubeClipper等开源项目，另外也在筹备发布OpenHydra等更多的开源项目，进一步推动开源软件在国内发展。

● 建木项目，2022年3月入选开源GitOps产业联盟生态图景，同年11月正式入驻GitLink引擎模块，2022年12月正式捐献于木兰开源社区。期间获颁“2022东北亚优秀开源项目”、中国开源云联盟“2022年中国优秀开源项目”、OSCHINA最火热中国开源项目社区、2022年中国开源创新大赛二等奖等。

- Skyline项目，2021年9月捐献于OpenInfra基金会，同年12月项目从孵化毕业，正式升级为官方（Official）模块，成为OpenStack标准组件的一部分。2022年10月，Skyline第一个正式版本随OpenStack Zed正式发布。
- OpenV2X项目，自2022年5月成立至今发展迅速，已发布Albany（2022年7月）、Beihai（2022年11月）及Columbia（2023年6月）三个版本，项目代码量累计为525,868行（Columbia版本）。2023年8月正式成为开放原子开源基金会孵化筹备项目。
- KubeClipper项目，2022年8月正式开源，2023年6月加入CNCF Sandbox。
九州未来获得OpenCloudOS社区2023年度突出贡献企业奖；九州未来高级技术总监黄舒泉作为黄金成员代表加入OpenInfra亚洲首届咨询委员会。

在过去十年、从A到Z共23个版本的OpenStack，九州未来总体贡献（Commit）排名全球第十一、全国第二，并在多个重要项目模块中保持贡献全球领先，为OpenStack贡献中国力量。

开源建木项目本年度推出v2.7版本；通过腾讯云计算TencentOS Server 2兼容性认证，共创国产操作系统新篇章；与ChatGLM团队及青田创新赋能中心共同搭建了一套ChatGLM-6B对话体验环境。

OpenV2X项目已经实现CSAE 53-2020合作式智能运输系统车用通信系统应用层及应用数据交互标准第一阶段Day I和第二阶段Day II定义的V2X典型应用场景7个，计划未来2-3年内通过行业伙伴开源协作方式是实现29个典型应用场景全部覆盖，成为国家新四跨协议的标准实现。

（7）蚂蚁集团

2021年6月1日，蚂蚁集团自研数据库OceanBase宣布正式开源，并成立OceanBase开源社区，社区官网同步上线。开源仓库包括OceanBase数据库内核、驱动以及相关的平台工具产品（例如开发者平台ODC、运维工具平台OCP-ExPress等），组织累计收获GitHub10000+ Star）。

该开源项目曾获得2023 OSCAR尖峰开源项目、尖峰开源企业（开源运营与生态建设）、“可信开源社区”评估证书。

蚂蚁集团参与贡献最多的开源项目包括：Flink/Flink CDC、SeaTunnel、DolphinScheduler、Bytebase、KubeBlocks。

OceanBase项目在2023年共推出正式版本2个，分别是OceanBase 4.1版本、OceanBase4.2版本。在

GitHub上Star增长到7250, Fork增长到1485、Issues增长到1226、PR增长到577。

截止目前, 社区拥有超300位贡献者, 超500家客户将OceanBase社区版应用于实际的业务生产系统。技术生态方面, OceanBase持续构建以"合作伙伴"为中心的产品技术文化, 广泛已经与多基础设施、数据集成、数据治理、应用集成服务商围绕OceanBase进行产品适配和对接, 有超过750个主流产品已加入OceanBase自研技术生态。2020年, OceanBase正式商业化, 客户数已超过1000家, 客户数年增长150%, 其中30%客户将其应用于核心系统。

(8) 平凯星辰

平凯星辰于2022年加入开放原子开源基金会, 成为白银捐赠人。TiDB开源社区是由平凯星辰(平凯星辰)发起成立的数据库开源社区, 是TiDB生态中的开发者、用户、合作伙伴共同分享和学习的平台, 在全球数据库领域已具有广泛的影响力。TiDB开源社区覆盖45个国家和地区, 累计超过2154位开源贡献者。TiDB线上社区asktug.com汇聚了3.3万名TiDB资深用户, 已经积累23,000+个问题帖, 90%的问题都得到了解决, 累计总回复243,000+个。

平凯星辰为数据库开发人才培养制作了开源数据库实操平台——Talent Plan, 受到了全球数据库技术爱好者(尤其是在校学生)的喜爱。作为开源项目, Talent Plan在GitHub上获得Star数超过9.8K。截至目前, Talent Plan已经被来自于250多所高校的2000多名学生学习使用。2023年, 公司成功举办第22届中国计算机学会中国数据库暑期学校, 来自清华大学、北京大学等29余所高校117名学员参加学习, 同时向中国计算机学会捐赠3年合作暑期学校工程实践全部实验。

平凯星辰发起的或作为主体维护的开源项目

截至目前, 公司累计发起超过220个开源项目, 其中包括分布式数据库TiDB、键值数据库TiKV、分析引擎TiFlash、混沌工程工具Chaos Mesh、开源分析工具OSS Insight、Talent Plan等全球知名开源项目。

TiDB: TiDB项目于2015年9月正式开源, 在GitHub上已总计获得超过36k颗星, 累计有1781位开源贡献者, 目前项目已经合并的PR数31594个, 已经解决的Issue数13133个。TiDB项目来自平凯星辰以外的社区贡献者占比 75 %以上, 吸引了腾讯云、韩国三星研究院等企业和机构的优秀开发者共同参与。

TiKV: TiKV项目于2016年4月正式开源, 在GitHub上已总计获得超过14.1k颗星, 累计超过620位开源

贡献者，项目已经合并的PR数11059个，已经解决的Issue数3594个。TiKV项目主要开发者来自平凯星辰、知乎、京东云、一点资讯等企业，来自平凯星辰以外的社区贡献者占比65%以上。

TiFlash: TiFlash项目于2022年4月正式开源，在GitHub上已总计获得927颗星，拥有贡献者超过90人。TiFlash项目也是支持TiDB数据库分析能力的核心模块。

ChaosMesh: Chaos Mesh于2019年12月正式开源，在GitHub上已总计获得超过6.1k颗星，累计超过207位开源贡献者，项目已经合并的PR数2515个，已经解决的Issue数1177个。

Talent Plan: 是公司数据库开发人才培养制作了开源数据库实操平台，受到了全球数据库技术爱好者的喜爱。作为开源项目，Talent Plan在GitHub上获得Star 超过9.8K。

2023年，TiDB项目荣获开放原子开源基金会“2023生态开源项目”。平凯星辰公司创始人黄东旭当选开放原子开源基金会理事。同时，公司入选权威非营利性国际组织BenchCouncil发布的世界首个开源贡献榜，排名中国企业第一，全球企业第十四，公司三位创始人刘奇、黄东旭、崔秋均入选“开源领域五十年人才榜”。

平凯星辰主要参与以下8个开源项目，并有代码反馈和相应贡献，包括：TiDB、TiKV、TiFlash、ChaosMesh、OSSInsight、Talent-Plan、RocksDB、Rust。

在2023年，TiDB共更新代码743万行，其中平凯星辰公司贡献368.88万行，占比49.6%；外部组织和个人贡献374.22万行，占比50.4%。全年，TiDB全球下载量约为28.1万，其中中国大陆14.3万，海外13.8万。全年，TiDB技术社区累计主题讨论发帖超过4700个，回复数超过50000个。在GitHub上，PR数量为901，issue数量为3.6K，Fork数量为5.7K，Star数量为36K。

(9) 思斐

思斐发起ShardingSphere开源项目，于2018年正式进入Apache软件基金会孵化器，2019年进入CNCF全景图，于2020年成为Apache顶级项目。2021年度Apache软件基金会年度报告中，ShardingSphere代码提交数量位列前十。

ShardingSphere 2023年斩获多项荣誉，2023年1月SphereEx斩获掘金引力榜「年度新锐企业」、「年度新锐人物」两项荣誉。在基金会中中国项目排行榜里，ShardingSphere名列第一。在中国项目OpenRank排名Top30榜单里，ShardingSphere位居第三。截至2023年5月，LibHunt发布的前3名Java Database开源项目榜单中，ShardingSphere始终名列前三。目前，借着领先的技术优势及广泛的业内

认可度，ShardingSphere也在持续促进众多不同领域与行业创新应用的落地，其覆盖的技术应用生态圈汇聚了超240000+名个人成员、534名核心贡献者以及225家合作企业。

截至2023年11月，ShardingSphere社区共吸引了全球19000+来自全球的开发者关注；产生了40800+技术创新碰撞（Commit），9200+共同技术探讨（Issue），18000+社区代码贡献（PR），200+开发者指南、性能白皮书、最佳实践等高质量社区深度原创内容。项目贡献者有555人，项目核心决策成员的身份包含海外成员，此外，成员的年龄分布于80后、90后、95后等，具有阶梯性及多样性。

（10）涛思数据

据“CSDN 2022中国开发者影响力年度评选”数据统计，在中国开发者个人贡献Top20中，有四位是来自涛思数据的开发者，在开源全球公司贡献Top50，涛思数据也名列其中。2023年，凭借在开源领域的突出贡献，涛思数据被开源中国评为“2023年度优秀开源团队”，并荣登了Bench Council（国际测试委员会）发布的多项开源相关的榜单，包括开源计算机系统成果榜、首个“开源贡献世纪榜（Open100）”、开源计算机系统机构榜等，深化了涛思数据在国际开源上的影响力。2023年5月，涛思数据以“银牌赞助商”的身份参与了国际开源数据库领域全球最具影响力峰会Percona Live 2023，同年10月，涛思数据以“金牌赞助商”身份参与了2023年度中国计算机大会，获得了众多开发者的关注。

涛思数据采用AGPL许可证，于2020年将TDengine的内核(存储和计算引擎和集群) 100%在GitHub开源，并尽最大努力打造开发者社区，旨在通过开源，快速获得市场反馈，完善产品和生态，吸引更多的开发者加入到项目中。发展5年时间后，TDengine在GitHub上的Star数已经达到了23.1k（统计至2024年8月16日），多次在GitHub全球趋势排行榜上排名第一。

TDengine在2019年开源后，很快发行了1.6版，2020年8月发行了2.0版，2022年8月发行了3.0版，以上版本全部进行了开源。

在众多开源的数据库中，TDengine的增长速度极其惊人。从2019年开源到2023年，其已成长为时序数据库领域Star数排名第三的新星项目，并在全部开源数据库产品中，位居第六；同时，保持着在开源时序数据库领域中最快的PR增长速度，截止到2024年3月28日，PR总次数已经达到20.4k。

相对于2.x版，3.0解决了业内的High Cardinality问题，能支持10亿条以上的时间线，100个节点以上的集群。它在分布式设计的基础上，实现了计算和存储分离，让TDengine从3.0起成为了一个云原生时

序数据库。而且3.0将流计算、数据订阅功能完全重构，极大地提高了效率，让TDengine为时序数据的处理提供了一个极简的平台。同时，TDengine 3.0重构了查询计算引擎，对SQL提供了更多支持，让很多复杂查询得以进行。

凭借着开源开放的力量，在产品和服务质量的双重保障下，TDengine的用户群体也在不断扩大，全球运行的TDengine实例数超过592.3k（统计至2024年8月16日），用户遍布全球。用户中既有使用开源版本的中小企业，也有很多全球500强企业。包括蔚来、理想、京东、顺丰等各行业客户，均通过使用TDengine，大幅提高了业务效率、降低了总拥有成本。在商业化方面，TDengine企业版及专业技术服务也得到很多客户认可，订单量呈现持续攀升态势。

为帮助众多的传统产业数字化转型，过去的一年，TDengine大力开发了PI System、MQTT、OPC等各种工业数据接口，支持与各种BI、AI工具对接，同时在阿里云、AWS、Azure、GCP上推出全托管的云服务TDengine Cloud。无需大量资金投入、无决策风险下，短时间内开发者就可以试用开通，并与AI应用、众多的分析工具、可视化工具集成，体验大数据以及AI的魅力。

涛思数据及所属人员获得的开源社区/项目荣誉（2023年度）

- 登开源计算机系统成果榜——Bench Council（国际测试委员会）
- 荣登首个“开源贡献世纪榜（Open100）”——Bench Council（国际测试委员会）
- 开源计算机系统机构榜——Bench Council（国际测试委员会）
- 2023年度优秀开源团队——开源中国
- 涛思数据创始人陶建辉荣登开源系统贡献榜——Bench Council（国际测试委员会）

（11）统信软件

Deepin（深度）社区于2004年发布国内首个基于Debian的Linux中文发行版，Deepin操作系统在2015年进入全球Linux发行版（Distrowatch）前十名，并于2015年加入Linux基金会，社区母公司统信软件是2023年开放原子开源基金会白金捐赠人。

目前Deepin社区注册用户近40万，社区论坛注册用户超过15万，是目前国内最活跃的桌面操作系统开源社区。国际影响力方面，deepen社区在全球有西班牙、巴西、土耳其、俄罗斯、斯诺伐克五大分

支社区，海外用户超300万，Deepin操作系统全球下载量超过9000万。

Deepin社区自身拥有软件仓库8100多个，提交PR超过16000个，Issues提交超过8700个。对于上游社区贡献，Deepin社区向Linux内核、GCC、Qt等上游社区提交补丁1340多个，Patch合并920多个，累计向上游社区提交代码超过13.7万行。

Deepin（深度）社区发起的或作为主体维护的开源项目

- Deepin: Deepin是一个开源的操作系统项目，拥有活跃的开源社区。
- 玲珑：“玲珑”是统信软件自研的开源软件包格式。它用于与dpkg、rpm等包管理系统共同实现应用包管理、分发、沙箱、集成开发工具等功能，目标是解决Linux长久存在的依赖问题和安全问题。玲珑项目已捐赠到开放原子开源基金会。
- Deepin-ide: Deepin-IDE即Deepin-unioncode，是由深度公司自主研发的具有多语言、跨平台兼容特性的轻量级集成开发环境，其开源许可协议为GPL-3.0。

Deepin（深度）社区及所属人员获得的开源社区/项目荣誉（2023年度）

- 中国开源创新大赛二等奖和优秀奖；
- 中国科学院软件研究所-“2023开源创新榜”优秀开源项目奖；（Deepin）
- 中国科学院软件研究所-“2023开源创新榜”优秀开源人物；（张磊）
- 开放原子开源基金会-2023生态开源项目（Deepin）
- 开放原子开源基金会-2023快速成长开源项目（DDE桌面组件）；
- “第六届世界声博会暨2023科大讯飞全球1024开发者节”年度创新产品奖；
- 开源中国-2023年度优秀开源技术团队奖；
- 入选2023年度上合国家软件产业国际合作优秀案例；

Deepin（深度）社区发起或捐献的最具影响力的开源项目年度进展

Deepin: 已脱离Debian上游，目前Deepin V23基于Linux内核直接构建，且在国内桌面操作系统发行版中率先支持的Wayland显示协议。累计收到了180名外部开发者，超过2700个代码贡献。

- DDE: 深受国内外喜爱的Linux系统桌面环境, 全球下载量累计超过9000万次。
- 玲珑: 全新的现代化软件包格式, 支持多种Linux发行版, 可以跨系统版本、跨不同发行版运行应用程序。软件生态已达1300+。
- Deepin-ide: Deepin推出的自研IDE程序, 支持DTK自研开发库, 目前已支持AI插件。
- DTK: Deepin系统和DDE桌面环境的核心开发库, 广泛用于 Deepin自研GUI应用程序中。

(12) 微众银行

2017年, 微众银行牵头研发的区块链底层平台FISCO BCOS完全开源, 并推动成立深圳市金融区块链发展促进会(简称“金联盟”), FISCO BCOS逐步成为中国联盟链技术领域的标准。

2019年, 微众银行成为Linux黄金会员并将自研的全球第一个联邦学习工业级开源框架FATE捐赠给了Linux基金会。

2021年, 微众银行先后向Apache软件基金会捐赠了插件式云原生事件基础设施Apache EventMesh和大数据计算中间件Apache Linkis两个开源项目。两个项目于2023年初先后顺利毕业成为顶级项目, 其中Apache Linkis是全球首个由银行机构主导、捐赠并毕业的Apache顶级项目。

2024年初, 微众银行将自研的开源项目CyclOne(一站式资源全周期管控平台)捐献给开放原子开源基金会openEuler社区, 成为国内第一家向openEuler项目群捐赠开源项目的金融机构。

过去几年, 微众银行陆续在人工智能、区块链、云计算、大数据等多个开源领域取得优异成绩, 截至2024年初, 微众银行对外开源项目已达36个, 这些项目受到了全球开发者的广泛认可, 获得了超过40,000个Stars和15,000多个Forks。

2023年度主要参与了OpenFunction、openEuler社区的贡献。微众银行中间件团队与EventMesh社区一起



完成了与FaaS平台OpenFunction社区的集成贡献并完成合并，重新设计了连接器（Connector）的架构，社区基于此架构开发了超过10多个常见组件的event connector（source/sink）插件，为双方项目吸引了更多的贡献者的参与，正在设计开发的Function架构将进一步完善EventMesh的Serverless化生态。同时，主导EventMesh项目加入了开放原子开源基金会的openEuler项目社区，基本完成了openEuler系统集成EventMesh的安装包，为云原生应用程序的开发和部署提供了操作系统的原生接入，用户可以更便捷的使用事件驱动基础设施服务。

微众银行发起或捐献的最具影响力的开源项目年度进展

FATE (Federated AI Technology Enabler)是微众银行AI部门发起的开源项目，为联邦学习生态系统提供了可靠的安全计算框架，并且建立了全球范围内规模较大的隐私计算联邦学习开源社区，自2019年2月开源发布以来，FATE累积发布40多个版本，覆盖40多种算法组件，GitHub代码仓库中Stars数量超5400个，微众银行累计提交超13000 Commits，超过23万行代码。

2023年推出了FATE 2.0版本，从应用层互联，调度层互联，通信层互联，异构计算层互联这4个方面实现多层次隐私计算平台间互联互通，助力数据要素市场开放互联。目前FATE开源社区的参与者多达1300余家企业和500余所高校和科研机构。

FISCO BCOS是微众银行牵头研发、全面开源、国产安全可控的金融级联盟链底层平台，在性能、易用性、安全性等方面具备行业领先优势，目前已迭代至第三代，并且全面通过电子标准院、中国信息通信研究院区块链系统功能和性能测试。

2023年，第三代FISCO BCOS共计发布更新9个版本，继续推进与WeDPR多方大数据隐私计算平台的融合，引入多种安全多方计算协议解决数据协作中的隐私保护问题，实现多方数据的安全可信协作，释放数据生产力。8月，FISCO BCOS架构论文还被国际顶会Supercomputing收录，获得了国际学术界的高度认可。依托超10万名成员、5000多家企业参与者，FISCO BCOS开源社区已发展成为最大最活跃的国产开源联盟链生态圈，涌现出400多个产业数字化标杆案例，覆盖政务、跨境、金融、医疗等多个行业领域，并从中筛选了19个领域、252个案例编订成册，正式发布《2023 FISCO BCOS产业应用发展报告》。

Apache EventMesh项目是由微众银行中间件团队主导，于2023年成功从Apache孵化器毕业成为Apache TLP项目中的一员。EventMesh作为一个动态的插件式云原生事件基础设施，提供了灵活、可靠和快速的事件分发与处理能力，通过信息传递，让不同的“应用和服务”可以更好的对

接、集成。Apache EventMesh进入Apache后总计发布12个版本，2023年新增了3位PMC成员、9位Committer和23位contributor，累计贡献者已超过340位。

2023年，在代码方面，Apache EventMesh新增494个Issue、819个PR、2575个Commits，进行了3次版本升级，分别是v1.8.0、v1.9.0和v1.10.0，其中新增了Eventmesh c-sdk、Eventmesh connector、Eventmesh filter/transformer以及 Eventmesh 支持k8s部署4大特性。在企业应用落地方面，新增了吉利汽车、中化能源科技公司、中国移动信息技术有限公司等，构建了多元丰富的开源开发者社区。同时，项目社区始终坚持共建的理念，向Openfunction社区、openEuler社区贡献了集成EventMesh的PR等。

Apache Linkis是微众银行大数据团队自主研发的大数据计算中间件项目，自2019年开源至今，已成为业界颇具影响力的大数据领域开源项目，并逐步形成广泛的应用生态圈。目前，开源社群用户总数超7600人，沙箱累计试用公司超2600家，收到超过110家企业已投入生产的反馈，生产环境支撑的数据量超400PB，生产服务的用户超5000人，涉及金融、电信、制造、互联网等多个行业。

2023年1月，Apache软件基金会官方宣布 Apache Linkis 顺利毕业，成为 Apache 顶级项目(TLP)。在项目版本迭代方面，2023年度Apache Link进行了4次版本迭代，涵盖9大新增特性功能，6项已实现的功能的优化增强。与此同时，Apache Linkis在2023年度积极与开源社区联动，同时注重与高校的合作。在社区运营的各项指标方面也均有提升，新增了2位PMC、3位Committer和25位Contributor，累计贡献者达到156人，在代码方面，新增了453个Issue，522个PR、775个Commits。

DataSphere Studio（简称 DSS）是微众银行自研的数据应用开发管理集成框架，基于插拔式的集成框架设计，及计算中间件 Linkis，可轻松接入上层各种数据应用系统，让数据开发变得简洁又易用。借助于 Linkis 计算中间件的连接、复用与简化能力，DSS天生便具备了金融级高并发、高可用、多租户隔离和资源管控等执行与调度能力。

2023年发布的DataSphereStudio1.1.2进行了11项特性的改进，将微服务进行合并，并且支持了微服务DSS-Server、DSS-Apps-Server多活部署与Apache Linkis 最新版本1.4.0完成集成。2023年度各项运营指标稳步提升，新增5位Contributors，累计贡献者达48人，123个Issue、20个PR、363个Commits。，同时积极参与开源生态的布道活动。

(13) 中兴通讯

中兴通讯在LF AI & DATA基金会孵化推理侧工具链Adlik，目前已经吸引累计86位全球贡献者，十余家贡献单位。Adlik项目经权威机构评测，推理侧加速性能处于全球领先水平。

中兴在ONAP前身Open-O贡献超过50%的代码。2017年成为ONAP初创会员，并担任微服务总线MSB、告警分析Holmes两个子项目的PTL。

中兴通讯支持中国运营商在Linux网络基金会发起的5G/6G电信能力平台xGVela的立项，并贡献Telco PaaS初稿。中兴通讯多年来也在开源开发者论坛和开源峰会积极发声，大力推动5G网络切片、网络管理自动化、网络智能化等开源与标准在5G中的结合与应用。

中兴终端对Mojo（AI编程语言）官方文档进行了中文翻译，已在GitHub/Vercel发布了国内首家Mojo文档（中文版）。

2023年中兴通讯开源专家孟伟竞选成为LF AI & Data董事会主席，这也是当年Linux基金会下设各董事会主席中唯一的中国企业代表。

中兴通讯及所属人员获得的开源社区/项目荣誉（2023年度）

- Adlik项目获得“中国开源云联盟优秀开源项目奖”，“CSDN中国开发者影响力项目奖”及“中国开源云联盟优秀开源项目”等三项奖；
- 获得开放群岛开源社区2022-2023年度优秀共建单位；
- 开源治理入选中国信通院《“源生万态”——中国通信行业开源创新发展案例集》；
- 孟伟荣获COPU2023年度中国开源优秀人物奖；
- 操作系统团队获得2023龙蜥社区卓越贡献奖；

中兴通讯参与以下10个开源社区，并有代码反馈和相应贡献，列表如下

参与开源社区情况				
序号	开源社区名称	单位所承担角色	单位加入时间	单位贡献情况
1	Open Infrastructure Foundation	黄金会员	2013	5,252 commits
2	Cloud Native Computing Foundation	黄金会员	2017	3,020 commits

3	Ceph Foundation	高级会员	2018	2,079 commits
4	DPDK Project	黄金会员	2017	732 commits
5	RISC-V国际协会(前身:RISC-V基金会)	高级会员	2018	NA
6	Linux基金会 AI&数据基金会	高级会员	2018	Adlik: 75K LoC + 322 commits
7	Linux基金会 网络基金会	铂金会员	2014	29,700 commits
8	Linux Foundation	白银会员	2015	39,700 commits
9	龙蜥社区	副理事长单位	2023	1,700+ commits
10	OpenCloudOS	副理事长单位	2022	600+commits

(14) 小米

小米作为长期活跃于Apache软件基金会的贡献者，深度参与并积极贡献于包括但不限于HBase、NuttX、Spark、Flink在内的15个由Apache软件基金会所孵化与管理的开源项目。比如，自2017年起，活跃在Apache NuttX社区，坚持开源共建，共拥有三名PMC，4名Committer，连续两年贡献社区超一半的代码修改，是社区最大的贡献者。21年小米加入LVGL社区，目前拥有2名Committer，去年贡献了社区34%的修改，也是LVGL开源社区最大的三方贡献者。除上述外，小米还在非常多的开源项目里做贡献，包括小米联合Intel共同开源的Wasmizer-TS编译器项目，以及OpenAMP，WAMR，Zephyr等。

小米及所属人员获得的开源社区/项目荣誉（2023年度）

- 小米“新一代Kaldi项目”获得开放原子开源基金会“2023快速成长开源项目”奖；
- 员工肖翔凭借其对Apache软件基金会的贡献及其在操作系统领域内的技术突破，荣获“2023中国开源优秀人物”奖。
- 2023年6月，江海、柴华、曹娟在2023开放原子全球开源峰会荣获“2023年度开源贡献之星”。
- 2024年3月，“Xiaomi Vela”获得“最具影响力IoT操作系统”奖项；
- 2024年3月，李桂丁凭借在NuttX社区的卓越贡献，获得2024开源创新先锋奖。

(15) 易捷行云

易捷行云从2014年开始参与多个国内外开源社区，并鼓励全员参与工作相关的各类开源社区，公司赞助了OpenInfra, CNCF, containerd, openEuler等等社区，荣获过相关社区的董事成员。积极参与相关开源社区的贡献，在多个社区获取重要的角色，包括Kubernetes、OpenStack等多个项目的核心评审员。

积极参与国内的开源事业，曾赞助开放原子开源基金会，参与开源创新委员会，参与openEuler、OpencloudOS等社区合作及共建。积极参与国内各类开源相关活动，包括开源创新大赛、OIF days China, 开源年会、kubernetes等等。

EasyStack 2023年聚焦公司主营业务，重点参与与公司业务相关的重点开源项目，包括OIF社区的Nova、Cinder、Neutron、Kata-container。CNCF社区的kubernetes, containerd, istio等项目。Linux社区kernel, OVN等项目。2023年，EasyStack张彬彬荣获OIF社区最佳新人奖。

EasyStack公司将开源写入公司的愿景，鼓励研发工程师参与各类开源项目，近年来在各个主流国内外社区都贡献突出。

在OpenStack社区的Nova、Cinder、Neutron、Keystone等核心项目贡献排名Top5，总计贡献Commits数318，修复bug数8个，贡献代码量12000+行，工作人天超过1000人天，主导其中heat项目PTL，katacontainer核心贡献者1人，主导国内、东南亚等区域活动，提升社区影响力。

在Ceph社区贡献次数119次，代码量6000+行，核心Committer 1名，作为赞助商协助推动开源社区透明性建设，推进社区规划及财务公示，官方网站关注运营模型介绍等信息。

在CNCF社区投入kubernetes、containerd、etcd, kube_OVN等多个项目，其中社区多个版本国内贡献TOP10，贡献核心需求10+，bug数数十个，是社区赞助商之一。

(16) 睿赛德

RT-Thread目前是全球GitHub Star数排名第一的开源嵌入式操作系统项目，已捐赠予开放原子开源基金会，同时RT-Thread也加入RISC-V基金会，支持国产RISC-V芯片的软件生态。

睿赛德主要在维护的开源项目是RT-Thread。生态构建进程：项目团队一方面持续优化内核性能，另一方面着力构建丰富的软件生态。通过与多家知名芯片厂商、云服务商及应用开发商深度合作，RT-Thread已集成众多实用的软件包，涵盖网络协议栈、文件系统、图形界面、AI算法等多个层面，极大

地简化了开发者的工作流程，缩短了产品开发周期。国际化进程：随着RT-Thread在国内市场的稳固地位，项目也开始放眼全球，积极拓展海外市场。不仅在国际开源社区中声名鹊起，入选GitHub年度最活跃项目Top10，还成功打入海外各大主流嵌入式芯片平台，成为了全球开发者青睐的实时操作系统之一。

2023年RT-Thread开发者大会在上海临港中心成功举办，RT-Thread社区技术负责人郭占鑫介绍了RT-Thread近期的技术更新以及版本迭代情况。其中，主要更新包括用户态方面的POSIX API开发、xmake构建工程、动态链接和静态链接、GNU APPs/LVGL/Weston等应用的开发，以及内核态方面的一键完成Smart、Nano版本切换、更强悍虚拟文件系统（支持vnode、mmap）、更好的多核及64位支持、硬件MPU、设备树解析、llvm工具链、pico-libc等。

2024年，RT-Thread 5.0.0版本发布，该版本增加了RT-Thread Smart特性，支持用户模式应用；除此之外，还为增加了原子特性，对调度器文件进行功能拆分；在组件层面还新增tmpfs文件系统，增加musl libc支持并完善了POSIX的支持；在bsp层面，新增了约40+个bsp等。

5.1.2 国有企业

(1) 中国移动研究院

2014年，中国移动以创始会员身份加入Linux基金会旗下OPNFV社区成为白金会员。

2016年，中国移动联合华为等公司在Linux基金会发起成立Open-O社区。

2017年，中国移动联合AT&T宣布全球最大的网络编排器开源社区ONAP正式启动。

2017年，中国移动代表中国公司首次荣获OpenStack开源社区Super User超级大奖。

2018年，中国移动、AT&T、DT、DOCOMO、Orange五家运营商联合成立O-RAN联盟。

2018年起中国移动成为Linux网络基金会白金会员，2021年至2023年，中国移动首席科学家冯俊兰博士担任该基金会董事会主席。

2020年，中国移动在Linux网络基金会发起成立全球首个电信云原生项目XGVela。

2022年，中国移动在OpenInfra基金会发起成立全球首个算力网络开源项目CFN (Computing Force Network)

中国移动发起的顶级开源项目包括OPEN-O、ONAP、XGVela、G-SRv6；联合其他公司发起的顶级项目包括OPNFV、O-RAN SC、Akraino等。

2023年，5个项目入选中国通信行业开源案例集。其中赵奇慧荣获OSCAR开源先锋人物奖。王升荣获OpenInfra Community Hero。郁志强入选中国开源先锋33人和Linux基金会开源学园开源推广大使等荣誉称号。

（2）中国联通

中国联通目前参与的开源社区主要涉及云原生、云存储、SDN控制器、白盒交换机、网络编排器等领域，包括OpenStack基金会（黄金会员）、Linux基金会（白银会员）、OpenEuler社区。在Linux基金会白银会员基础上，中国联通作为白银会员/高级会员参加了LFN、SODA、CNCf子基金会。

中国联通基于沃云孵化的自研海量数据存储项目YIG已在LF SODA开源，实现了中国联通开源社区牵头项目“零”的突破，该项目已在乐视、联通、紫光云大范围应用。

此外，中国联通深度参与openEuler社区开源建设，积极贡献两大核心技术产品FastBlock和SafeGuard。

FastBlock是新一代高性能全闪块存储产品，重点面向信创数据库场景和中小企业虚拟化场景。FastBlock通过分布式技术实现系统高可用、高可靠，在仅采用普通NVMe介质情况下，并发时延性能超过开源Ceph 5倍以上。SafeGuard是基于eBPF和KRSI的Linux安全防护系统，通过系统漏洞检查、弱配置检查、入侵检测、病毒扫描等功能，实现对文件、进程、网络等操作的拦截及审计记录，达到事前主动防御、事后审计追溯的效果，全面保障操作系统的安全性和稳定性。目前SafeGuard已在联通云CULinux中广泛应用，为联通打造高安全操作系统做出重要贡献。

5.2 开源基金会发展现状

5.2.1 开放原子开源基金会

开源已成为支撑关键基础软件开发迭代、引领新一代信息技术群体突破的主流创新模式。为确保集众智、汇众力的开源模式充分发挥出赋能千行百业创新的潜力。开放原子开源基金会（简称基金会）应运而生。2020年6月，基金会在北京正式成立，在民政部注册登记，是我国唯一一个国家级开源基

基金会，补齐了我国开源生态发展的核心环节。发起单位包括阿里巴巴、华为、浪潮等在内的10家科技龙头企业。基金会具备独特的中立属性，在增强各方互信、凝聚各方资源、聚合广泛生态、保障开源项目长周期发展等方面具有明显优势。

基金会成立后，广泛借鉴全球领先的开源基金会发展经验，构造了较为全面的开源基金会治理架构和运营体系，基金会运营的重点项目有序发展、基础设施能力初步具备、品牌影响不断扩大、支撑保障能力不断强化。

在拓展开源项目方面。基金会开源项目全部来自于企业捐赠，23个项目进入正式孵化，还有28个项目处于孵化筹备期，涉及开源芯片、基础软件、汽车软件、工业软件、人工智能、云原生、数据库等重点领域。明星项目社区主要有OpenHarmony（华为捐赠）、openEuler（华为捐赠）、OpenTenBase（腾讯捐赠）等。其中，截止2024年6月，OpenHarmony吸引超7800名核心贡献者，累计产出超1.1亿行代码，230家厂家的608款产品通过兼容性测评；openEuler社区开发者达1.8万名，成员单位超1500家，用户超243万人；OpenTenBase社区委员会正式成立，招商银行、英特尔等首批11家成员企业加入。

在建设代码托管平台方面。截止2024年6月，基金会牵头建设的AtomGit代码托管平台即将正式交付，平台注册开发者已达到100万。线下协作空间、兼容性实验室、数据中心、开源项目展厅等硬件设施建设基本完成。此外，针对部分地方相关个性化需求，初步实现了为地区定制化提供开源项目成长、监测、统计等系列功能。

在提升基金会品牌影响力方面。基金会以“两会一赛”为依托，建立活动体系，加强开源文化推广，提升基金会品牌影响力。一是举办生态大会和开发者大会，在北京、江苏等地筹备并成功举办开放原子开源生态大会（全球开源峰会）2次以及开放原子开发者大会1次，政府领导、专家学者、厂商代表、顶尖开发者积极参与，加速了开源文化传播。二是完成第一届开源大赛，在江苏省工信厅、无锡市工信局等大力支持下，设置50个赛题，发放奖金超过1330万元，吸引4980支队伍、4.34万名选手参赛，参赛选手来自1000余所高校、科研院所等，在基础软件、工业软件等领域解决191项技术难题，产生573项技术成果，其中31个获奖团队与地方签署意向落地协议，各方面成效远超同类型赛事。三是加强基金会自媒体矩阵建设及运营，在微信、微博、抖音、小红书、今日头条等开通基金会账号，形成品牌传播矩阵。

在培育开源人才方面。以开放原子“校源行”公益项目为抓手，推动开源人才培育，目前已覆盖200

余所高校，支持67所高校实现社团运营活动常态化。2024年以来，“校源行”完成线上线下系列活动90余场，拓展中科院软件所等6家培训合作伙伴，培训200余名开源导师，并推出OpenHarmony人才认证，即将推出openEuler和OpenTenBase认证。

在强化协议供给方面。审定发布MIT、Apache2.0、GPLv3等多个主流开源许可证中译文，出版《商业开源》译书，发布多份国内外开源司法判决书译文，围绕人工智能模型、开源硬件等领域，开展新一代许可证研究与编制工作。

在加强国际交流方面。目前，基金会已与全球Linux基金会、Eclipse基金会、Fossasia社区等开源组织建立了较为紧密的联系。例如，开放原子开源基金会与Eclipse基金会就OpenHarmony项目群战略合作签订的合作协议，这是全球开源历史上第一次两个基金会通过代码、品牌、认证等方式共同发展一个开源生态，为开源业内提供了合作的新典范。

5.2.2 天工开物开源基金会

天工开物开源基金会（全国第二个、省级第一个开源基金会）于2023年3月获得重庆市政府批复成立，由清华大学，华中科技大学，重庆大学，中国信科，中科院等知名高校和企业联合发起，致力于推动中国开源事业发展的非营利性独立法人机构。

天工开物开源基金会本着“开放、务实、专业、创新”的发展理念，推动开源技术赋能国产工业软件和制造业数字化转型发展；遵循共建、共治、共享原则，系统性打造开源开放框架，搭建企业、高校开源社区，提升软件行业及应用单位的协作效率，赋能千行百业。

目前，基金会业务范围主要包括募集资金、专项资助宣传推广、教育培训、学术交流、校企合作、开源生态建设、咨询服务等业务。

2023年5月13日，以“天工开物，面向未来”为主题，由重庆市发展改革委、重庆市科技局、重庆市经济信息委西部科学城重庆高新区管委会、中国人工智能学会联合主办，天工开物开源基金会承办的开源未来发展峰会在重庆隆重召开。工开物开源基金会宣布了多项与开源社区的合作共建项目：

- 天工开物开源基金会与COPU共建开源平台
- 天工开物开源基金会与开源指南针项目合作共建开源评估生态
- Linux Foundation开源软件学园渝源社区落地

- 天工开物开源基金会与开源中国代码托管合作共建
- 天工开物开源基金会联合产学研用各机构发布了《开源战略合作重庆宣言》

2023年10月19日，以“探索开源新疆界，共创数字化未来”为主题的2023开源新疆界·天工开物多元合作峰会在武汉正式拉开帷幕。来自各个开源组织、开源社区的负责人，以及开源从业者、高校教师与企业界同仁共聚一堂，以开源为题，以大会为媒，展开了一场深刻的行业趋势大讨论。同期成立SigStore社区，社区将通过组织技术交流、项目孵化、人才培养等系列活动，鼓励开发者参与到开源项目中，激发创新的火花，同时，社区还将加强与国内外开源组织的合作，共同推动开源技术的研究和应用。

2023年10月19日，三大项目正式捐赠天工开物开源基金会

达坦科技——DistroSim项目

- 采用软件分布式计算技术加速的数字芯片验证框架；
- 采用高级软件编程语言Python、C++实现数字芯片验证；
- 应用于高性能网络RDMA的FPGA原型验证。

武汉简视科技有限公司——Voice ImPPrint TTS项目

- 基于VALL-E X模型开发的智能声音印记技术；
- 复刻原声音的音色、韵律、音调，成本更低、周期更短、效果相近；
- 为短视频、短剧出海，提供理想的译制配音解决方案。

上海交通大学重庆人工智能研究院——“微著”项目

- 用于解决多体问题的科学计算平台；
- 算法原创、软件自研、硬件加速、国产适配、平台开放、代码开源；
- 基于原创算法，在110个大规模问题中，43%的年均算法效率提升大于年均硬件能力提升；约13%的问题，算法提升大于硬件提升效率的十倍。

5.3 其他开源组织

5.3.1 COPU

2004年7月22日成立于北京的COPU (China OSS PRomotion Union, COPU) 由致力于开源软件文化、技术、产业、教学、应用支撑的企业、社区、客户、大专院校、科研院所、行业协会、支撑机构等组织自愿组成的、民主议事的民间行业联合体, 非独立社团法人组织。COPU的宗旨是为推动中国开源软件 (Linux/OSS) 的发展和应用而努力; 为促进中国与全球关于开源运动 (Linux/OSS) 的沟通、交流与合作而努力; 为促进全球开源运动 (Linux/OSS) 做出贡献而努力。截止目前, 已经连续主办了18届“开源中国开源世界高峰论坛”和16届“中日韩三国IT局长OSS会议暨东北亚开源软件推进论坛”, 成为最具国际影响力的中国开源组织之一。

自2015年始, COPU提出发展基于开源的人工智能。

2018年Linux基金会组织开发、推广跨平台的深度信息技术 (包括大数据, 云原生, 区块链, 人工智能等), COPU随即提出发展基于开源的深度信息技术。2023年, COPU更是提出为适应人工智能时代需要, 应向更开放、更透明方向修改传统的开源技术, 发展基于开源的人工智能。

5.3.2 北京开源芯片研究院

北京开源芯片研究院 (以下简称“开芯院”) 在 2022年 8月正式加入RISC-V国际基金会 (以下简称“RVI”), 成为高级会员拥有理事会席位和技术指导委员会席位。开芯院首席科学家包云岗担任理事会席位具有投票权, 固定参加每双周进行的理事会会议, 负责决策整体战略、发展路线、统筹资源、推动标准化及促进RISC-V技术的商业采用方面发挥重要作用。同时在技术委员会参加审查和批准技术规范、推动开源软件工具和文档的开发、组织技术讨论和决策等。

开芯院是2023年RISC-V中国峰会主办方, 该峰会是RISC-V全球三大峰会之一。开芯院于 2023 年 10 月正式加入RISC-V 软件生态系统 (RISE) (The RISC-V Software Ecosystem (RISE)) 项目是国际上唯一认可的服务于 RISC-V 软件生态的组织。由Linux Foundation Europe (LFEU) 和 RISC-V 基金会共同推动成立。开芯院是技术指导委员会成员。开芯院系统软件组基于香山项目发展结合社区在虚拟化和UEFI 固件的开发计划进行开发, 现已补丁合并, 修复 UEFI 在 RISC-V 平台上启动DxeCore 阶段的 hang 问题, 在RISC-V 平台上开发UEFI 的开发者可以直接使用。修复开发者在使用 Gmac 驱动的 hang 问题。

“香山”高性能开源 RISC-V 处理器项目，是目前国际上性能最高的开源高性能 RISC-V 处理器核，在开源芯片技术体系中最接近 L3 级别的开源处理器核项目，作为国际上最受关注的开源硬件项目之一，已在全球最大的开源项目托管平台GitHub上获得超过4300多个星标，形成超过581个分支。

5.3.3 北京智源人工智能研究院

北京智源人工智能研究院（简称“智源”）是一家专注于人工智能领域的研究机构。

智源研究院在人工智能的多个方面开展前沿探索，目前下设多个研究中心，包括生命模拟研究中心、健康计算研究中心等。其研究方向涵盖大模型、生命模拟、智能计算等领域。

2024年6月14日—15日，“2024北京智源大会”在中关村展示中心盛大召开。该大会被誉为“AI界春晚”，今年的大会上国产大模型成为主角，百度、月之暗面、智谱AI、零一万物、面壁智能等大模型公司备受关注。

智源在推动人工智能发展方面有诸多举措，例如发布北京人工智能数据运营平台并开源超大规模高质量数据集；联合多所高校推出首个多任务长视频评测基准MLVU；与Hugging Face联合推出开放中文大语言模型榜单-旗鉴榜；联合多家机构推出自动化多样性信息检索评测基准AIR-Bench等。

同时，智源也注重人才培养和社区建设，通过智源学者、智源大会、智源社区、青源会等项目，汇聚了顶尖学者和青年才俊，建设了活跃的AI生态圈。

第六章 开源国际化发展现状

6.1 国际开源基金会和企业对中国的贡献

6.1.1 Linux基金会

Linux基金会（The Linux Foundation，简称LF）搭建了一个中立且值得信赖的平台，供全球开发者和企业共同开发、管理和扩展开源项目。目前，中国在开源代码和项目的贡献上正以迅猛的速度增长，已经成为一个极为重要的参与者。以云原生计算基金会（CNCF）为例，根据CNCF云原生计算基金会CTO Chris Aniszczyk介绍，中国已经成为第二大开源项目贡献国。截至目前，中国贡献了Harbor、Volcano、Dragonfly等32个项目，大约占CNCF所有项目的四分之一。

在过去20年里，Linux基金会已经从单一支持Linux内核项目扩展到支持众多不同的项目社区。目前，有超过2300名成员和数十万名开发者参与到一些最为关键和活跃的开源项目中，这些项目在云计算、安全、区块链和Web等跨行业的技术领域展开合作。

到了2021年，Linux基金会在它所支持的数百个开放技术社区中的贡献和参与度持续上升。特别是中国区域的成员增长极为显著，仅CNCF中国区成员的增长就达到了80%。

随着对专业人才需求的不断增长，供应商中立的培训和认证变得日益重要。在中国，已有超过5500人获得了Kubernetes认证管理员的资格。2021年，Linux基金会向全球，包括中国的个人，提供了免费的培训和认证奖学金，数百人通过与非营利组织的合作获得了这一荣誉。

除了上述贡献外，Linux基金会还通过举办各类活动和建立合作，为中国的开源发展注入了强劲动力。

- 2007年，Linux基金会组织欧美资深开源开发者来北京支持COPU召开的“Linux开发者国际会议”，与国内开发者展开深入交流。
- 2008年，作为主要助力之一，Linux基金会助力COPU无偿聘任数十名国际开源领袖、开源大师为智囊团高级顾问，从而建立了中国的开源高地，也是创新高地、科技高地和人才高地。
- 2017年，Linux基金会成立一批旗下跨界的基金会（如云原生、区块链、人工智能等）并建立孵化

器，为中国和全球培养人才与发展技术作出贡献。

Linux基金会每年都会参加由COPU举办的"开源中国开源世界高峰论坛"，在这个年度盛会上，它作为圆桌会议的核心团队之一，深入探讨国际开源技术的最前沿发展、国内外开源生态发展中遇到的挑战及其解决方案。通过这些深入的交流和讨论，Linux基金会致力于推动开源社区的全球合作与发展。

此外，Linux基金会针对不同生态领域成立了众多新基金会，并见证了越来越多的中国企业加入其合作行列。例如，2021年上半年，RISC-V基金会的成员数量翻了一番，其中许多新成员来自中国。同时，LF AI & DATA、LF Networking、O3DF、OpenSSF、CNCf等基金会的董事会中，中国企业的代表也在显著增加。由中国企业发起的开源基金会，如NextArch，正在全球范围内迅速崛起并蓬勃发展。

在2023年里，Linux基金会积极在中国举办各类开源活动，为中国开源生态系统的成长注入了强劲动力。

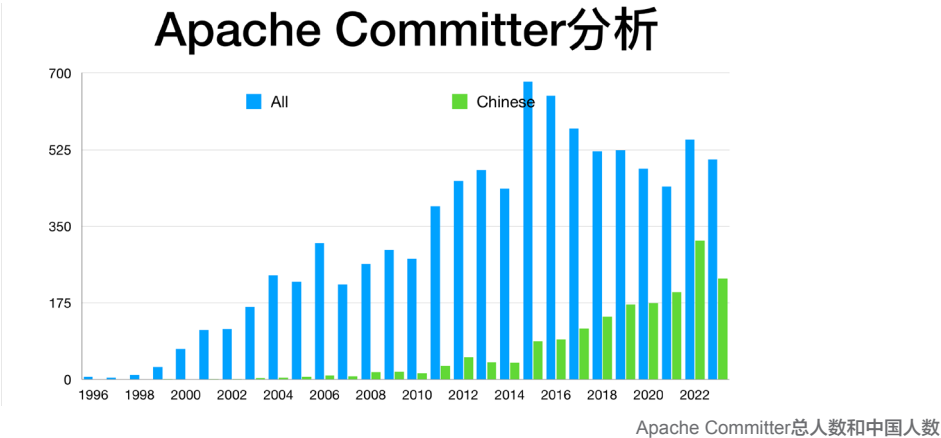
- 3月，Linux基金会在香港举办了亚太领导力峰会；
- 5月27日，在2023全球开源技术峰会（Global Open-source Technology Conference, GOTC）上，与上海浦东软件园签约，成立Linux基金会亚太区开源社区服务中心；
- 6月，LFOSSA成功举办了RISC-V国际人才认证中国区发布会，并发布首个RISC-V认证RVFA的中文版；
- 8月，LFOSSA与RISC-V International联合举办了RISC-V中国峰会之国际人才暨开源技术与生态论坛；
- 9月，LFOSSA与CNCf在上海联合发布云原生人才培养计划；
- 10月，LFOSSA与中国20多所大学、开源社区及机构加入天工开物基金成立的开源创新教育联盟，一起推动开源教育的创新与发展。

未来，Linux基金会和其亚太团队将继续推动中国企业和社区参与全球开源合作，培养更多开源人才。

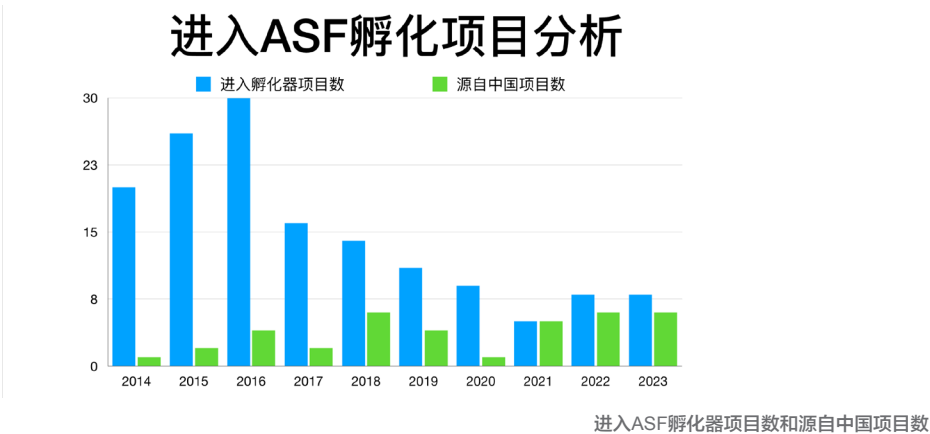
6.1.2 Apache软件基金会

2012年之前，国内与ASF（Apache Software Foundation，简称“ASF”）接触的公司主要集中在外企，例如Intel中国研发中心、IBM中国研究院、eBay中国等。但在2012年之后，华为、阿里巴巴、百度、京东、腾讯、蚂蚁等公司开始陆续将项目捐赠给ASF，在ASF的孵化下，多个项目成为顶级项目。

伴随着越来越多来自于中国的项目进入到ASF，中国的Committer数量也在不断增加。如下图所示，在2010之前，中国开发者参与度并不高。自2014年开始，中国Committer数量逐年递增，直至今日，中国Committer数量占到新增Committer一半左右的比例。



后图展示了2014年之后源自中国Apache孵化项目与所有进入到ASF的孵化项目对比，可以看到2021年之后来自于中国的孵化项目占比越来越大。



2023-2024年, Apache在中国举办的ASF相关会务活动

- 2023年7月, ALC-Xi'an在六朝古都西安正式成立。
- 2023年7月26日-28日, CommunityOverCode Asia在美丽的西子湖畔杭州召开线下大会。
- 2023年8月18日-20日, 由Apache软件基金会举办的官方全球系列大会CommunityOverCode Asia首次中国线下峰会于北京顺利举行。
- 2024年5月, ALC-HangZhou正式成立, 进一步链接杭州的阿里, 蚂蚁, 字节等大厂的开发者资源, 共谋开源发展大事。
- 2024年7月, ALC-Shanghai正式成立, 为华东地区的开源开发者提供了一个连接, 交流, 成长的好环境。

6.1.3 开放基础设施基金会

OpenInfra基金会专注于打造一个多元化的开源社区, 并推动开源基础设施软件在实际生产中的广泛应用。在全球187个国家/地区的110,000多名社区成员的共同支持下, 该基金会不仅托管了一系列开源项目, 还积极开展社区实践活动, 覆盖了人工智能、容器云原生应用、边缘计算和数据中心云等多个领域。在中国, OpenInfra基金会旗下的OpenStack、Kata Containers项目以及算力网络工作组尤其受到社区成员的广泛欢迎, 并呈现出活跃的增长态势。

(1) OpenStack社区

2022年, 中国的OpenInfra社区发展迅速。凭借众多的志愿者和庞大的开源基础设施组织基础, 中国仍然是世界第二大OpenStack社区。在过去的一年了, OpenStack技术架构委员会提交、接受并发布了两个来自中国贡献的新项目: Venus project提供日志管理服务的一站式解决方案; Skyline项目由UI和UE优化的现代OpenStack仪表盘。

Venus项目由浪潮数据发起, 回馈社区。这有助于发布全文/多维检索、典型错误分析、调用链分析等功能。

Skyline项目最初由九州云发起并贡献, 现在有超过四个组织参与上游开发。近期, OpenStack韩国用

户小组也开始了参与Skyline项目的预研和评估工作。

除了新近发布的项目，在过去的一年里，已经与OpenStack社区中开发成长了许多年的Cyborg继续展示其发展以及更多元化的上游社区。在上一版中，全球有超过25个组织参与了Cyborg上游开发，其中包括一众来自中国的企业与组织，如浪潮集团、麒麟软件。

（2）Kata Containers社区

Kata Containers在2022年取得了重要进展，尤其是在中国社区。社区已完成Kata Containers 3.2.0版本的发布，增加了机密容器对GPU设备的支持，以及对设备管理的抽象。

中国联通和浪潮数据发布了两个新的用户案例研究，针对Kata容器的大规模部署；蚂蚁集团的Kata Container使用白皮书也在2022年发布，在全球获得了很高的知名度，甚至获得了超级用户奖；300多位开发者在微信群里积极交流，讨论技术问题，分享部署。

2023年3月份，Kata Containers中国社区也在成都举办了疫情恢复后的第一次线下技术聚会，来自社区的30多位开发人员参加了聚会并分享各自的实践案例以及技术研发。据统计，Kata Containers上游社区去年共有来自18个组织的140多位开发者提交了1363项修改，其中将近50多位开发者来自中国，包括一些企业、科研学术机构和大学同学。

（3）算力网络（CFN）工作组

2022年，新的算力网络（CFN）工作组成立，最初由中国移动发起。该小组旨在利用相关技术制定应对共同挑战的解决方案，推动技术落地，逐步构建下一代开源基础设施技术。工作组得到了广泛认可，2022年共有16个组织加入工作组。

此外，2023年12月1日，OpenInfra Days China 2023在北京海航万豪酒店盛大启幕。时隔两年，OpenInfra Days China重返线下，聚集了全球社区的技术专家、行业领袖以及国内开源基础设施技术企业的代表，围绕开源基础设施项目的最新进展、生产实践以及所面临的挑战进行面对面分享与交流。

6.1.4 云原生计算基金会

云原生计算基金会全称Cloud Native Computing Foundation，简称CNCF。CNCF中国区蓬勃发展。企业会员超过90家，Kubernetes官方认可服务供应商（KCSP）44家，Kubernetes官方培训伙

伴（KTP）14家。截至2023年，CNCF（云原生计算基金会）在全球已经吸纳了超过173个项目，中国已成为第二大开源项目贡献国，来自中国的代码占比从2016年的3.8%增长到2022年的13.2%。目前有超过60个中国企业、团队成为了CNCF会员。

2023年，CNCF依旧注重云原生人才培养成功设计并最终确定了36门崭新课程。在此基础上，推出了20门全新课程，包括6门edX、MOOC和14门线上课程。中国区合作伙伴共进行超过300场CKA、CKAD、CKS等Kubernetes线上线下面授培训，培训人数已超过万人。

CNCF推出4个新的Linux基金会官方认证：

- ICA（Istio Certified Associate）
- HFCP（Hyperledger Fabric Certified Practitioner）
- FCFP（FINOS Certified FDC3 Practitioner）
- Skillcred -SC106（Pipelines with Tekton）

另外，以下认证已在Beta测试阶段，预计在2024年登场。

- KCSA（Kubernetes and Cloud Native Security Associate）
- CCA（Certified Cilium Associate）
- CAPA（Certified Argo Project Associate）
- CGOA（Certified GitOps Associate）
- FSOSD（FINOS Financial Services Open Source Developer）

6.1.5 RISC-V国际基金会

2011年5月，加州大学伯克利分校研究团队发布了一套全新的开放指令集RISC-V。2015年，汇聚全球100多家单位的非营利组织RISC-V基金会正式成立。

为降低潜在的技术出口法律限制风险，减轻来自地缘政治的潜在破坏，同时为确保基金会更加中立、

开放和包容，2020年，RISC-V国际基金会将注册地迁移至瑞士，此举受到全球RISC-V爱好者的广泛好评。在过去一年，RISC-V国际基金会在中国有如下贡献：

- 2023年5月26日，RISC-V国际基金会首席执行官Calista Redmond女士、首席技术官Mark Himelstein先生等来华参与“RISC-V开源处理器芯片生态发展论坛”与在场嘉宾分享了其推动开源处理器芯片生态发展的努力与贡献。
- 2023年8月23日，2023 RISC-V中国峰会在北京盛大开幕。本届峰会由北京开源芯片研究院主办，中关村创业大街科技服务有限公司、广州市智能软件产业研究院承办。会议吸引了来自北京大学、清华大学、中国科学院、科技企业、投资机构等2000人参会。
- 2023年8月31日，RISC-V国际基金会CEO Calista Redmond女士担任特聘顾问应邀出席“RISC-V工作委员会成立大会”，RISC-V工委由中国电子工业标准化技术协会牵头，联合国内38家重点企业共同组建。
- 2023在Linux基金会的协助下推出了RVFA（RISC-V Foundational Associate）的中文版。

6.1.6 IBM

近年来IBM在AI，云原生，可持续计算，边缘计算，量子计算等前沿技术领域采用开源、开放的模式面向全球进行探索与创新。

在IBM 2024 Think大会在开源创新，AI和自动化，生态合作三条主线上均有重磅发布和创新进展。IBM同时将最新的人工智能平台watsonx及其背后的开源技术和其他AI相关开源技术一并介绍给中国。

今年，中国开源软件推进联盟与IBM联合发布《可信赖的企业级生成式AI白皮书》，旨在为中国企业提供借鉴，助力企业基于自身的业务需求与数据来构建开放、可信、具有业务针对性的企业级AI应用与模型。

IBM在Hugging Face平台上开源了Granite和Merlinite两个大模型的同时携手Red Hat开源了instruct lab项目，这个项目实现了Large-Scale Alignment for ChatBots技术，使得模型训练所使用数据的公开透明。我们通过IBM watsonx和该技术显著改进了Granite模型。在有效提高了模型可解释性的同时

缓解了GPT-4等专有LLM生成合成数据的合法性风险。

数据是AI生态的重要组成部分，2023年4月，IBM收购了Ahana，成为湖仓查询引擎Presto Foundation的成员，并在整个社区中成为仅次于Meta的最大贡献者。IBM在Presto的多个领域均有显著贡献，尤其是在最新发布的Presto C++中，IBM是主要贡献者。Presto C++的性能比Presto Java提升了2-3倍。

随着AI和大模型应用上云，云计算领域开始出现一些新的热点领域，如可持续计算，机密计算，云原生可观测性等，IBM也和国内外CNCf社区的伙伴们有着广泛的合作。

GPU能耗在数据中心不断增加，为了更好地为数据中心进行能耗规划在内的诸多挑战，CNCf sustainable TAG（云原生可持续技术指导工作组）正式成立。IBM同Intel，Red Hat合作，承办了CNCf sustainable TAG 2023年在国内的CNCf sustainable week（云原生可持续计算周）活动。上千人在线观看了技术分享，也收到了来自国际社区对于活动结果的认可和祝贺。在2024年推进CNCf sustainable TAG，决定将在亚洲时区召开月度工作会议，开始筹备2024年的云原生可持续计算周活动。

在机密容器（Confidential Containers）开源社区，IBM与社区的伙伴密切合作，进一步将零信任安全原则纳入机密容器框架，增强了安全性和易用性。IBM还公开了机密容器在不同行业中的各种用例。例如，在医疗保健领域保护患者数据。在金融行业，确保敏感的金融信息在机密容器内处理，帮助实现合规并防止欺诈。

聚焦云原生可观测性，IBM也在不遗余力的将可观测性领域的最佳实践经验通过各种形式回馈分享给开源社区，尤其是Instana团队在公开博客中介绍了通过将可观测性技术和生成式人工智能相结合以实现将治理融入AI生命周期中监控管理的相关实践。

IBM在系统软件特别是编译器领域与业内建立了紧密的联系，与国内龙头企业华为，阿里平头哥，字节，RISCv社区等都有各种形式的交流与探讨。IBM中国编译器团队自2010年成立以来，深耕编译器后端优化技术，取得了多项成果，在业内两大开源编译器领域GCC和LLVM，林科文和陈正先后被社区任命为maintainer。从2022年开始，积极投入Rust语言开源社区，为IBM硬件平台带来Rust开发语言支持的同时，也为社区贡献了非常多的质量和生态提升的代码。与此同时，以传统编译器为基础，在其他系统软件的开源领域诸如高性能科学计算库OpenBLAS，调试器LLDB，JIT Linker以及仿真器QEMU都有涉足及贡献。

IBM在2021年成立了RISC-V的兴趣小组，来促进围绕着RISC-V开源处理器的创新，协作和生态的建设。IBM参与Open Horizon开源边缘计算和Kepler可持续开源社区的研发人员，开发了Open Horizon和Kepler对RISC-V的支持，从而把RISC-V的使用生态扩展到云和边，以及可持续的应用场景。

在量子计算领域，经过六年开源努力，汇集成千上万用户的持续热情和数百名贡献者的持续支持才得以实现的，Qiskit 1.0版本发布是一个重要的里程碑。这个版本的发布标志着Qiskit开发进入了一个新阶段，该阶段以性能、稳定性和可用性为中心。

IBM非常重视开源协作，从量子 and 区块链到容器、人工智能和操作系统，积极领导当今最具影响力的项目，并创建新项目以推动明天的技术发展。

6.2 中国企业对国际开源基金会的贡献

时至今日，尽管全球开源界普遍认可中国企业和开发者对全球开源产业的巨大贡献，但仍有人持中国仅是全球开源资源的索取者和使用者的成见，这是一种误解，更是偏见。

事实上，在20世纪90年代中期，中国开源还处于启蒙阶段，国内最早接触这一领域的技术先驱者尝试积极参与国际开源社区的代码共享活动，虽然是凤毛麟角，但星星之火终成燎原之势。

21世纪初，随着国内Linux系统和LAMP技术栈的兴起，中国开源进入萌芽阶段，一些全球知名企业在华研究机构成建制地组织中国开发者，系统地参与到国际知名开源社区中。国内开源社区如雨后春笋般层出不穷，开源文化在当时那一批20岁出头的中国开发者心中埋下了种子。

2008年，作为中国开源发展史上的一座重要里程碑，见证了Linux基金会、GNOME基金会、Apache软件基金会以及OpenOffice基金会在北京首次举办技术峰会，标志着它们在亚洲的首次亮相。这一年，中国开源贡献者的地位获得了全球开源界的广泛认可，标志着中国开源进入了新的发展阶段。

在过去13年中，中国开源开发者开始习惯于用流利的英语和全球开源人士分享他们的技术经验。根据COPU统计，在2016年至2017年在华召开的开源国际会议上发表的论文中，来自国内和国外的的论文比例为4:6，即有40%的开源论文来自中国开发者。

中国的开源力量正在以惊人的速度迎头赶上，在开源开发者贡献度、开源社区活跃度、开源基金会参与度等方面，我们均已取得不俗的成绩。这些来自中国的开发者和科技企业正在向全世界呈现全新的中国印象，中国逐步实现了开源世界的大国担当。

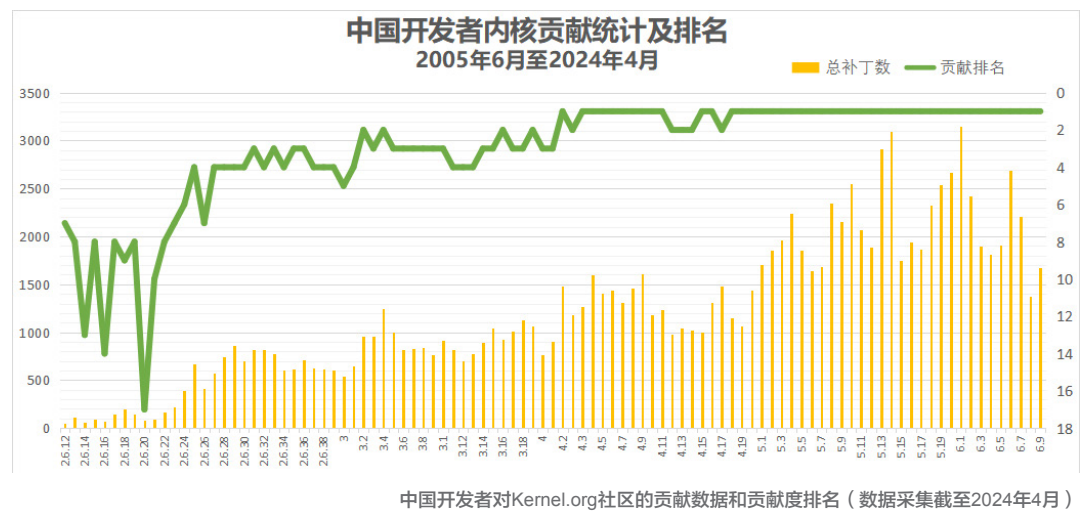
6.2.1 对Linux基金会的贡献

Linux基金会董事会成员共23人，其中中国成员2人，占比为8.6%。分别为华为的侯培新与腾讯的单志豪。

Linux基金会目前共有中国会员139家，相较于2020年的70家增涨了98.6%。其中白金会员2家（约占17%）；黄金会员2家（约占18%）；白银会员78家（约占7%）；Associate（准会员）23家（约占6%）。

同时，中国开发者已经成为Kernel.org社区最大贡献群体。

在Linux基金会，甚至全球开源生态中，Kernel.org社区的象征意义无可替代，开源界的创世一代至今依然活跃在这个社区，因而Kernel.org至今仍常被作为一个开源指数来度量一个企业或一个区域的开源繁荣和发展程度。为更加全面地展示中国在开源领域的成长过程，我们分别通过中国开发者和中国科技企业对Kernel.org社区的补丁贡献数据来呈现。

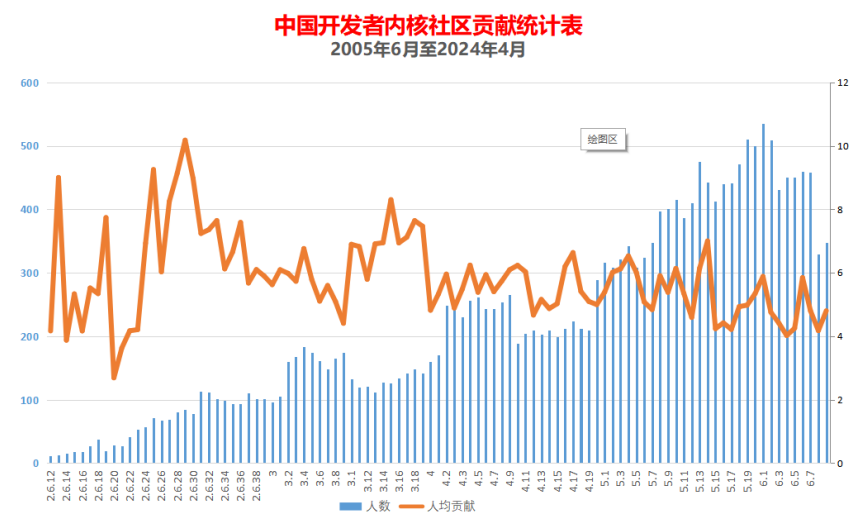


为了更直观地呈现内核社区中中国开发者的参与度和贡献度的变化，我们节选了部分版本的贡献度进行对比，从下表可看出，在过去的19年间，参与Kernel.org社区的中国开发者的规模提升了40倍，对社区的贡献绝对数量提升了50倍，对Kernel.org的贡献排名近五年来保持世界第一。

内核版本	中国贡献补丁数量	中国贡献补丁占比	中国开发者数量	国家贡献排名
V2.6.12	50	2.90%	12	第七名
V3.0	540	5.90%	96	第五名
V4.0	771	7.45%	160	第三名
V5.0	1444	11.27%	289	第一名
V6.0	2670	17.34%	500	第一名
V6.9	1671	12.72%	348	第一名

部分版本贡献度对比（数据采集截至2024年4月10日）

在统计过程中，我们还有一些发现。通过对比19年来中国开发者数量的变化，可以推算出中国开发者对Kernel.org社区的人均贡献度，其整体变化如下图所示。



中国开发者对Kernel.org社区的人均贡献度（数据采集截至2024年4月）

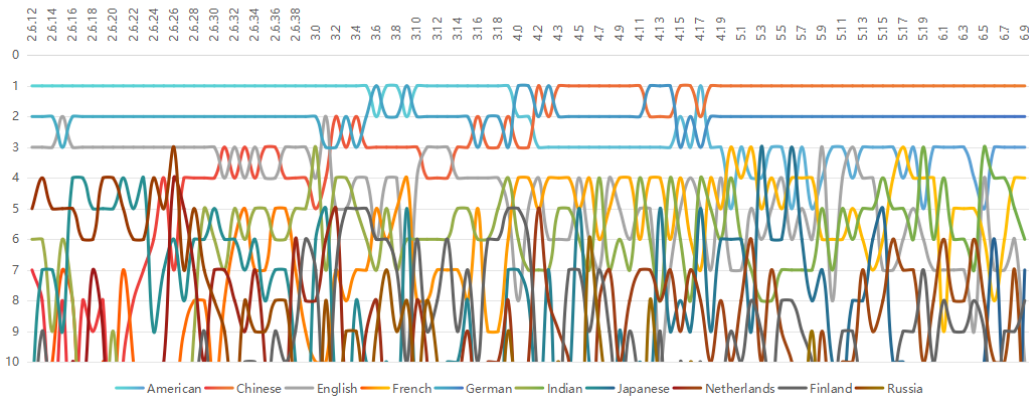
通过对比最近三个内核版本排名前三国家的开发者人数和人均贡献值，可以发现现今中国开发者的人均贡献度远小于德国和美国，进而我们可以推断出未来中国开发者的人均贡献度和绝对贡献数量还有很大提升空间。

国家	V5.0 贡献排名	V5.0 贡献人数	V5.0 人均贡献	V6.0 贡献排名	V6.0 贡献人数	V6.0 人均贡献	V6.9 贡献排名	V6.9 贡献人数	V6.9 人均贡献
中国	第一	289	4.99	第一	500	5.34	第一	348	4.8
德国	第二	72	13.62	第二	68	14.55	第二	73	17.93
美国	第五	47	8.80	第三	33	14.48	第三	32	16.71

Kernel.org社区三个内核版本排名前三国家的开发者人数和人均贡献值

纵观Kernel.org社区总体的开发者贡献，我们会发现中国、美国、德国、英国、法国、印度是该社区的主要贡献群体，其中中国的发展速度远远超过上述其他国家。

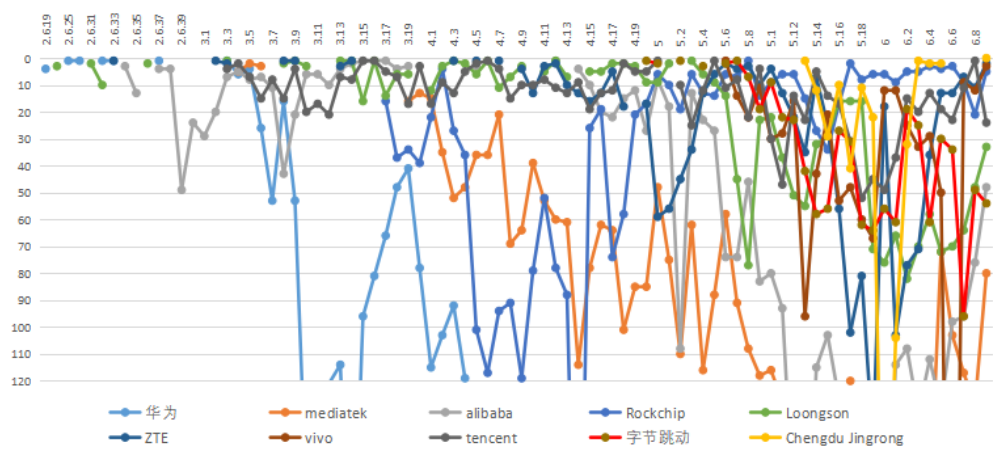
主要国家开发者内核社区贡献排名变化
2005年6月至2024年1月



主要国家开发者在Kernel.org社区的贡献排名（数据采集截至2024年4月）

从另一个角度，我们同样可以发现中国企业在Kernel.org社区中的贡献度越来越大，相关排名持续上升，下图是2005年6月至2024年4月的内核社区贡献排名中国前十企业的动态变化，包括华为、联发科、阿里、瑞芯微、龙芯、中兴、vivo、腾讯、字节跳动、成都蓉蓉。

十大中国企业内核社区贡献排名
2005年6月至2024年4月



十大中国企业在Kernel.org社区中的贡献排名（数据截止至2024年4月）

6.2.2 对Apache软件基金会的贡献

截至2023年6月，Apache软件基金会源自中国的活跃开源项目共34个，其中有17个项目成为了顶级项目。分别是：

序号	已有源自中国的ASF顶级项目（17个）	新增源自中国的ASF顶级项目（6个）
1	Apache Doris（2018/07/18-2022/6/15）	Apache SeaTunnel（2021/12/09 - 2023/05/17）
2	Apache InLong（原TubeMQ，2019/11/03-2022/6/15）	Apache EventMesh（2021/02/18 - 2023/03/22）
3	Apache DolphinScheduler（2019/8/29-2021/04/08）	Apache brpc（2018/11/13 - 2022/12/22）
4	Apache ECharts（2018/1/18-2020/12/16）	Apache Kyuubi（2021/06/21 - 2022/12/22）
5	Apache Ozone（2018/11-22-2020/10/21）	Apache Linkis（2021/08/02 - 2022/12/22）
6	Apache IoTDB（2018/11/18-2020/9/17）	Apache ShenYu（2021/05/03 - 2022/07/28）
7	Apache APISIX（2019/10/17-2020/07/15）	
8	Apache ShardingSphere（2018/11/10–2020/4/16）	
9	Apache Submarine（2019/10/16,Hadoop spin-off）	
10	Apache Dubbo（2018/2/16–2019/5/15）	
11	Apache Skywalking（2017/12/8–2019/4/17）	
12	Apache Griffin（2016/12/5–2018/11/21）	
13	Apache ServiceComb（2017/11/22–2018/10/17）	
14	Apache HAWQ（2015/9/4–2018/8/15）	
15	Apache RocketMQ（2016/11/21–2017/9/20）	
16	Apache CarbonData（2016/6/3–2017/4/19）	
17	Apache Kylin（2014/11/25-2015/11/18）	

ASF全球有800余位成员，目前已知并互动较为频繁的中国及华人成员有29位。其中，姜宁在2022年3月ASF举行的成员会议上被选举成为新任董事。Apache软件基金会现有中国企业赞助商6家，分别是白金赞助商：华为、腾讯云；黄金赞助商：百度；白银赞助商：阿里云、滴滴出行；银牌赞助商：小米。

6.2.3 对开放基础设施基金会的贡献

OIF基金会董事会15人，其中中国成员5人（占33.3%）。分别为浪潮张炳玲；九州云黄淑泉；中国移动李翔宇；蚂蚁集团王旭；华为周君毅。OIF现有中国成员中包括白金会员1家（占20%）；黄金会员6家（占46%）；白银会员3家（5%）；支持机构13家（占4%）。

白金会员	黄金会员	白银会员	支持机构	
华为	九州云	易捷行云	腾凌科技	迎广科技
	中国移动	腾讯云	技嘉科技	思特奇
	中国电信	长江计算	中标软件	中科曙光
	中国联通		数码港	Synology
	中兴通讯		中国电子技术标准化研究院	普崴存储信息技术公司
	新华三		银信科技	云宏
			嘉值信息	

开放基础设施基金会

6.2.4 对云原生计算基金会的贡献

据CNCF的统计，超过20%的开源项目来自中国，贡献度排名上升至世界第二。在云原生应用上，68%的机构在生产过程中使用容器，31%的开发者称所在单位的容器使用量已经超过5000个。CNCF理事会共29名成员，其中中国成员2人，分别是阿里云的易立、华为的任旭东以及火山引擎（字节跳动）的张鑫。CNCF现有TOC成员11人，其中中国成员1人是来自华为的王泽锋。

CNCF基金会目前共有中国会员84家（见下表），其中包括3个白金会员（占15.8%）、10个黄金会员（占37%）、35个白银会员（占5.9%）、1个研究院所（占25%），以及3个最终用户会员（占3.3%）。

白金会员	黄金会员	白银会员				研究所	最终用户
阿里云	蚂蚁集团	灵雀云	云信达	PingCAP	易捷行云	浙大软件工程实验室	HiTRUST
华为	百度	比格容器云	爱捷云	启明信息	平安科技		汇付天下
火山引擎	中国电子云	宝兰德	中国金信	青藤云安全	沃趣		vivo
	DaoCloud	博云	谱云科技	中创中间件	中科驭数		
	新华三	东方国信云	软通动力	斯普信	云轴科技		
	浪潮	铁科院	联想	探真科技			
	京东云	中国移动	漫道科技	时速云			
	腾讯	中国系统	默安科技	东方通			
	中心通讯	中国联通	领航动力	世纪互联			
	中国电信天翼云	小佑科技	网易数帆	睿云智合			

6.2.5 对RISC-V国际基金会的贡献

RISC-V基金会的日常运营由董事会负责，由4名常委和21名委员构成，其中的中方委员有五名，RISC-V基金会设立了技术委员会，成员共32人，其中包含11名中国成员：晶心科技的苏泓萌；阿里云的陈炜；希姆计算的范福杰;成为资本的沈海滨；腾讯蓬莱实验室的高剑林；中科院计算所的石侃；华为的Peter Chun；中兴通讯的石义军；天津飞腾的赵Tom，PLCT实验室的吴伟；北京开源芯片研究院的包云岗；RIOS实验室的谭章熹。

在RISC-V这一轮发展浪潮中，中国企业、高校和科研机构发挥了巨大作用。有数据显示，中国有超过400家企业关注、参与和使用RISC-V指令集进行开发。据统计，在RISC-V基金会中，成员分为三个层次（见下表），在基金会核心成员（Premier Member）中，中方成员12家，占比达到48%。在战略成员（Strategic Member）中，中方成员41家，占比达到24%，在社区成员（Community Organization Member）中，中方成员14家，占比达到8%。

核心成员	战略成员			社区成员
阿里云	安路科技	浪潮	群芯闪耀科技	中国石油大学
晶心科技	北京奕斯伟	联发科	算能	重庆大学产业技术研究院
成为资本	中科昊芯	擷发科技	声智科技	华东师范大学
华为	智芯公司	南京沁恒微	进迭时空	兰州大学
天津飞腾	北京微核芯	芯来科技	智成电子	国立清华大学
RIOS	群芯微	PerfXLab	兆松科技	OpenEuler社区
希姆计算	深度数智	睿思芯科	上海瓶钵	山东大学
腾讯	青岛本原微电子	瑞芯微电子	优矽科技	上海交通大学
中兴通讯	兆易创新	睿赛德科技	超睿科技	台湾RISC-V联盟
北京开源芯片研究院	高云半导体	摩联科技	紫光展锐	北京微芯区块链与边缘计算研究院
中科院计算所	乐鑫科技	赛昉科技	芯原微电子	COMPASS实验室
中科院软件所	先楫半导体	万向区块链	创景科技	广东省智能家电创新中心
	北京君正	中科蓝讯	芯昇科技	广东省新一代通信与网络创新研究院
	广东跃昉	云豹智能		中科院深圳先进技术研究院

RISC-V理事会中方成员

在2018年12月RISC-V基金会首次宣布要迁至瑞士后，中国企业界和学术界加速了参与和融入RISC-V基金会的进程，其中中国RISC-V产业联盟和中国开放指令生态（RISC-V）联盟发挥了巨大的促进作用。自2020年开始，中国国内的RISC-V生态快速完善，并开始涌现越来越多的RISC-V产业化案例。

众所周知，中国一直在芯片领域处于受制于人的窘境，整个产业迫切走出一条自力更生、自主创新之路。RISC-V由于开源的特性，自诞生起就在中国产业界备受瞩目，大量中国芯片公司正在成为RISC-V的拥趸，加大对RISC-V的投入。而中国具备的电子制造业生态优势，也成为RISC-V应用前景最乐观的地区，RISC-V在中国的市场容量和发展潜力巨大。这些因素成就了双方“互利互助、合作共赢”的天作之合。

另外，RISC-V生态中的相应软件、工具链还有待完善。RISC-V指令集开源并不意味着CPU核心也同时免费授权，这对芯片公司提出了很高的设计和研发要求。客观上，目前RISC-V只是帮中国芯片产业走上快车道，但前路漫漫，还需要我们持续努力。在未来的RISC-V生态建设中，中国企业需要平心静气、戒骄戒躁，避免在资本的推动下过度炒作概念，同时加强全球产业合作，避免产业生态碎片化趋势。

6.2.6 对其他国际开源基金会的贡献

目前OIN社区现有中国成员共121个。

OIN社区中国成员				
华为	长城汽车	快手科技	一加	诺瓦泰克
腾讯	广州汽车	上海微创医疗	嘀嗒	巨杉软件
小米	浪潮	天珑移动	哔哩哔哩	天网
上汽集团	蚂蚁集团	宇视科技	深度	希云
阿里巴巴	浙江吉利	紫光展锐	炬芯科技	易捷思达科技
百度公司	中国银联	JOYY	一铭软件	新浪微博
海尔集团	上汽通用	大疆创新	麒麟软件	广州汇宏资讯
北汽福田	浙江大华	紫光国芯微	达闼	河南鸿祺
比亚迪	迈瑞医疗	魅族科技	布罗德林克	普华基础软件
海康威视	美团	东软	龙芯中科	iAuto

奇虎360	字节跳动	中科曙光	紫光云	安东
伟易达	商汤	TCL	萤石网络	香港建筑学院
金山云	易软天创	京东	泰坦发动机	上海和联电子
iASPEC	上海飞爱	山东巨奇数码	上海DS科技	联合堆栈
iASPEC	优帆科技	红旗	链接Omnia	上海科盛
安太信息	安徽开源	上海百星	PimFrame	上海东格智
上海FLUX	上海嘉天	上海坤旭网络	上海超星	黑泽明操作系统
中信知识产权	西瑟	上海乐杰电子	AllMobilize	幽灵霓虹灯
SORC	CRD-HK	ipfsbit/星际比特	香港开源	上海骏美特
上海创米	开源社	加密货币操作系统	CNX软件	V-checker
达信通成	厦门GME	HardenedVault	麒麟信安	Soundmax
		深圳市酷亚科技		

6.3 COPU中沙开源合作进展及成果

COPU在促进中国沙特开源合作方面的进展和成果。从背景、COPU的合作策略、已取得的工作成果、下一步计划等方面展开，详细阐述了COPU如何通过一系列具体行动，推动中沙两国在开源领域的深入合作。

6.3.1 背景与战略意义

近年来，在全球科技竞争日益激烈和地缘政治格局不断变化的背景下，中国与阿拉伯国家之间的合作交流变得愈发重要。作为"一带一路"倡议的重要组成部分，中国与沙特阿拉伯的开源合作正在迎来新的发展机遇。

2022年12月，中沙两国元首签署了《中华人民共和国和沙特阿拉伯王国全面战略伙伴关系协议》，同时两国政府签署了《中华人民共和国政府和沙特阿拉伯王国政府关于共建"一带一路"倡议与"2030愿景"对接实施方案》。这标志着中沙合作进入了一个新的历史阶段。在这一背景下，中国方面表示愿意与沙特在包括数字经济在内的多个领域展开全面合作。与此同时，沙特也迫切需要中国的科技和数字化技术来助力其"2030愿景"的实现。

沙特阿拉伯正在积极推进其“2030愿景”计划，其中数字化转型是核心内容之一。沙特政府高度重视信息技术的发展，特别是在人工智能、云计算、大数据等领域。在这一背景下，中国的开源技术和数字化经验对沙特具有重要的战略价值。正如沙特阿拉伯王国的通信和信息技术大臣阿卜杜拉·斯瓦哈在2023年7月接受《南华早报》记者采访时表示，“中国信息技术实现了世界级转型，沙特渴望复制成功故事。中国在过去几十年的数字化转型中积累了丰富的经验，特别是在开源软件和开放技术标准方面取得了显著成果。通过开源合作，中国可以帮助沙特加速其数字化进程，同时也为中国企业开拓新的市场机会。

基于这一战略背景，COPU常务副秘书长谭中意联合OpenTEKr 发起人狄安自2023年初开始，积极开展了与沙特的一系列交流对话工作。他们与沙特的相关政府部门、社会组织、高校科研机构以及科技企业建立了广泛的沟通渠道和合作关系。这些机构和组织包括沙特阿拉伯通信与信息技术部（MCIT）及其下属的数字政府管理局（DGA）和数字企业家创新中心（CODE）、沙特人工智能局（SDAIA）、沙特开发者协会（Parmg）以及沙特国王大学等。

中沙开源合作不仅有助于推动两国技术创新和数字经济发展，还能促进文化交流和人才培养。通过开源项目的共同开发和维护，两国的技术专家和开发者可以建立长期的合作关系，推动知识共享和技术传播。

6.3.2 COPU的战略定位与实施方针

COPU在推动中沙开源合作中采取了明确的战略定位和实施方针。以“促进源于中国的开源技术在沙特以共建的方式落地”为目标；以开源先行，商业跟进的方式推进合作为战略。具体实施方针为：

- 与沙特当地的开源组织建立战略合作关系；
- 在人工智能、操作系统等开源技术领域与沙特合作，成立联合工作组和联合项目；
- 通过企业深度合作的模式，推动合作研发成果的商业化落地；

6.3.3 已取得工作成果

在COPU和OpenTEKr 的密切合作下，由谭中意和狄安组成的“开源老DJ”团队积极推动，中沙开源合作交流已经取得了显著成效：

(1) 成功组织三期中方IT企业代表团访问沙特

第一期（2023年11月）：COPU组织了包括开源中国、浪潮数字企业、统信软件、万里数据库、飞轮科技、道客等企业的代表团访问沙特。代表团与沙特通信与信息技术部（MCIT）、数字政府管理局（DGA）、数字企业家创新中心（CODE）等机构进行了深入交流。

第二期（2024年3月）：COPU组织了包括蚂蚁集团在内的企业代表团参加沙特年度科技大会 Leap 2024。Leap大会是中东地区最大的科技盛会之一，吸引了全球数千名科技专家和企业代表。蚂蚁集团在大会上展示了其在区块链、人工智能和金融科技领域的最新成果，并与沙特本地企业进行了多场合作洽谈。

第三期（2024年5月）：COPU组织了包括华为、用友、明道云等IT企业的代表团参加沙特低代码、云和网络安全峰会并考察多个沙特机构和组织。

(2) 跟沙特开源组织Parmg签署战略合作备忘录

2024年2月21日，COPU常务副秘书长谭中意受邀参加沙特年度开源峰会Open Source Summit Saudi 2024，并发表主题演讲。会上，COPU与会议组织方即沙特著名开源组织沙特开发者协会（Parmg）正式签署了《开源战略合作备忘录》，共同推动沙特阿拉伯和中国在开源领域的交流与合作。

这份备忘录的签署标志着中沙两国在开源领域的合作进入了一个新的阶段，为双方在开源技术、人才培养和项目合作等方面的深入交流奠定了基础。

(3) 联合发起成立AI Desktop工作组

2024年6月，中沙AI Desktop Working Group（工作组）正式成立。该工作组由中国开源推进联盟（COPU）和沙特开发者协会（Parmg）共同发起，旨在推动新一代AI驱动的桌面操作系统的研发和应用。

工作组的成立成员包括来自中国和沙特的多家知名科技企业和研究机构，如统信软件、英特尔（Intel）、联想（Lenovo）和沙特著名IT企业Sinee。工作组计划在未来一年内开发出一套适合中国和中东市场的AI桌面操作系统原型，并在沙特进行试点应用。

这一举措不仅体现了中沙两国在高科技领域的深度合作，也为两国在AI和操作系统领域的技术创新提供了重要平台。

(4) 促进企业间实质性合作

通过COPU的牵线搭桥，多家中国IT企业已经与沙特方面开始了深入的合作洽谈。案例包括，国内某大型企业信息化软件研发企业与沙特某著名家族企业科技公司签署战略合作协议，共同拓展沙特和中东市场；另一家中国开源企业则与沙特某科技企业签署开源社区合作协议，先在开源社区项目上展开合作。

为了更好地服务国家战略，帮助中国开源企业出海，与沙特及阿拉伯世界共建基于开源生态的数字命运共同体，COPU计划从以下三个方面展开未来工作：

- 运营常态化：每周与沙特各方保持定期远程视频会议，维护合作关系；继续组织中国科技企业赴沙考察团，2024年下半年计划组织两次考察活动；开始组织沙特方代表来华考察，促进双向交流。
- 组织体系化：在COPU指导下，联合沙特相关单位探讨多种组织形式，包括共同发起成立开源创新合作中心等。
- 成果显性化：推动中国科技企业以开源生态共建的方式在沙特进行技术能力输出和落地；实现中沙之间基于开源的技术商业合作和市场共拓。

COPU在推动中沙开源合作方面已经取得了显著的进展。通过组织IT企业沙特考察团、跟沙特开源组织签署战略合作协议、联合成立技术工作组等一系列举措，COPU为中国开源技术在沙特的落地和发展铺平了道路。未来，COPU将继续深化与沙特在开源领域的合作，推动更多实质性的技术交流和商业合作，为构建中沙数字命运共同体做出积极贡献。

在全球科技竞争日益激烈的今天，开源合作不仅是技术创新的重要途径，更是促进国际交流、增进互信的重要纽带。COPU在中沙开源合作中的积极实践，为推动“一带一路”倡议在数字领域的深入实施提供了有力支撑，也为中国开源技术“走出去”探索了新的模式。我们有理由相信，随着合作的不断深入，中沙两国在开源领域的交流将会结出更加丰硕的成果，为两国乃至全球的数字经济发展注入新的活力。

附录 1：开源大事记

1991年

- 中国开源诞生元年，中美合作引进UNIX SVR4.2。陆首群、杨天行和张克治牵头，代表中方代表团与美国AT&T公司合作，美方将当时最新开发的UNIX版本——UNIX SVR4.2源代码向中方开放（除美国外，中国是全球获得源代码的第二家）。此后，中方组织全国软件专家、程序员等200多人共同翻译、编辑、出版了“UNIX SVR4.2中文版本”。中美合作引进UNIX SVR4.2”这一历史事件被Linux基金会收录，后被作为中国开源诞生的标志性事件，1991年也就此成为中国开源的诞生元年。

1992年

- 成立合资UNIX公司，产生深远影响。国内计算机软硬件厂商与USG合资成立UNIX（中国）系统技术有限公司，对后续UNIX/Linux在中国的发展和普及，以及开源文化的传播有深远影响。

1994年

- Linux与自由软件被引进中国。Linux操作系统和大量自由软件被宫敏引进中国，国内技术人员开始接触到Linux。同时，自由软件的火种开始在国内传播。

1997年

- 建立“中国自由软件库”。通过国家信息中心的帮助，宫敏在中国经济信息网上建成“中国自由软件库”。为国内技术人员了解、学习和使用Linux及相关自由/开源软件提供了便利。

1998年

- LVS（Linux Virtual Server，Linux虚拟服务器）诞生。用于实现负载均衡，其代码进入了Linux 2.4和2.6版本的官方内核中。章文嵩是这一开源软件创始人。
- 开源Linux图形用户界面支持系统“MiniGUI”进入开发。由魏永明发起，历经十余年的迭代更新，MiniGUI可支持Linux/uClinux、eCos、μC/OS-II、VxWorks、ThreadX、Nucleus、pSOS、OSE

等操作系统和数十种SoC芯片。

- 冲浪平台成功实现中文Linux操作系统第一次商业应用。该司研发的Xteam Linux中文操作系统成功在北京市政府研究室办公OA系统中实现应用。
- 中国第一批开源社区——阿卡社区（AKA）在清华大学成立。后该社区转变成为亚嵌公司，AKA早期的多数成员成为中国最早的一批Linux人才。
- 国内Linux内核工程师开始参与国际开源项目。正在自动化系读大二的学生吴峰光开始接触Linux，发现Linux的文件预读算法磁盘I/O性能较差，开始尝试对Linux内核进行改进。吴峰光后来成长为国内对Linux内核贡献最多的工程师。

1999年

- 中文版红旗Linux和RedOffice进入研发。由中国科学院软件所的孙玉芳牵头，基于Linux和OpenOffice进行了中文版红旗Linux和RedOffice的研制。
- 国内专业开发者社区CSDN成立。蒋涛创立中国软件开发者社区CSDN，目前已成为中国最大的开发者社区。22年来，CSDN为中国培养了几代开源开发者，同时也见证了中国开源快速发展的20年，目前正积极向第三代开源社区转型。
- 蓝点最早发布中文Linux版本——BluePoint Linux。蓝点团队发布汉化的BluePoint Linux，这是全球最早的中文Linux版本。同年，蓝点软件技术（深圳）有限公司创立。

2000年

- 共创软件联盟成立。在“863计划”支持下，国内高校、科研院所、IT企业联合发起成立“共创软件联盟”。联盟在开源许可证规则下对863计划的软件成果进行开源培育和孵化，提出以“开放源代码的协同创新模式”作为中国基础软件发展的主要路径，对我国开源产业发展起到了重要推动作用。
- 北京市科委筹建软件产业基地公共技术支撑体系。为提升北京软件产业工程化水平，北京市科委筹划建设北京软件产业基地公共技术支撑体系。

2001年

- Sun中国工程研究院成立。研究院内部设立Mozilla、Office、Java等小组，是外资公司首次在中国设立全职参与开源社区研发的研发中心，为中国培养了大批开源人才。
- Intel中国开源技术中心（OTC）成立。该中心于2003年正式更名为Intel开源技术中心。经过近20年发展，中心为中国开源产业界培养了大批顶尖专家。

2002年

- 国内第一个Linux社区发行版发布。黄建中以“CJacker”的ID在中国Linux公社发布了基于Redhat8再发布的Magic Linux版本，Magic Linux是国内第一个Linux社区发行版。
- 北京软件产品质量检测检验中心成立。负责运营管理公共技术支撑体系。

2003年

- IBM中国Linux解决方案合作中心在北京中关村软件园正式启动。该中心由IBM公司和北京市政府合作，联合众多国内软件开发公司，目的是为国内提供Linux实验平台，培养Linux软件开发人才。
- 中国科技论文在线网正式开通。作为中国最早的开放存取仓储之一，中国科技论文在线网为科研人员提供了方便、高效的交流平台，让最新科研成果得以及时推广，科研思想得到有效保护。

2004年

- 中、日、韩三国政府在北京签署“开放源代码软件合作备忘录”。这是三国首次将软件开源提升到国家级的战略合作层面。备忘录内容包括：三国政府将为开放源代码软件，尤其是Linux的开发和应用创造政策环境、加强Linux对三国语言的支持、推动Linux软件的研发和商品化以及在三国中的应用、大力推广Linux在电子政务/企业信息化和远程教育等领域中的应用等。7月，由三国政府倡导的Linux平台Asianux1.0发布。
- 中国科学院和国家自然科学基金委员会签署了《柏林宣言》。这一举动表明了中国科学界和科研资助机构支持开放获取的原则和立场。
- 中国开源软件推进联盟（COPU）成立。该联盟是一个非政府组织，由致力于开源文化、技术、产业、教学、应用的企业、社区、院校、科研院所等单位自愿组成。联盟的宗旨是推动中国开源软件的发展和应用；促进中日韩以及中国与全球开源运动的沟通、交流与合作；促进中国为全球开源运动做出贡献。

2005年

- 国家自然科学基金委员会正式发布“国家自然科学基金基础研究知识库”。该知识库收集国家自然科学基金资助项目成果的研究论文全文，向社会公众免费开放。
- Mozilla中国中心成立。Mozilla中国中心是由中国科学院软件研究所和Sun中国工程研究院共同发起，经Mozilla Foundation正式授权的非赢利性机构，该中心依托在中国科学院软件研究所。

2006年

- 中国开源软件推进联盟（COPU）主办“2006年第一届开源中国开源世界高峰论坛”。主题为“开放标准，开源架构，开源生态系统与应用解决方案”，大会讲师24位，包括担任COPU智囊团高级顾问的全球开源领袖及跨国企业专家，与会中外代表约250人。
- 首届中国Linux内核开发者大会（CLK）成功举办。Intel、清华大学、AKA社区共同发起第一届中国Linux内核开发者大会（CLK）。该大会延续至今，成为中国开源领域的顶级技术峰会，发掘出吴峰光、马涛等一批内核领域的专家。

2008年

- Linux（Kernel）、Apache、OpenOffice等国际知名开源社区在北京召开全球技术峰会。这是国际开源组织首次在亚洲国家举办技术峰会，是对中国开发者为社区所做贡献的肯定，也成为中国开源发展史上重要的里程碑。
- 阿里巴巴开始研发孵化Dubbo（开源分布式服务框架）。这是一款高性能、轻量级的开源Java RPC框架。2019年5月16日，Apache软件基金会董事会决议通过了Apache Dubbo的毕业申请，成为顶级项目。
- 开源爱好者红薯创建开源中国（OSCHINA）社区。该社区十多年来推动了中国开源领域的发展。
- 76位华人为Linux Kernel贡献822个补丁，2008北京Linux开发者研讨会在京召开。在COPU与国际Linux基金会合作举办的“2008 北京Linux开发者研讨会”上，中国新锐与国际开源社区资深大师同台研讨，在Linux Kernel 2.6.29版本11232个补丁中，由76位华人（占获得成功的志愿者总数6.85%）向社区贡献822个补丁（占全部补丁7.20%）。

2010年

- Linux基金会在中国设立分支机构。Linux基金会任命Cliff Miller出任基金会中国首席代表，这也是Linux基金会首次在中国设立分支机构，希望通过此举进一步推进Linux平台在中国的普及化。

2011年

- 阿里开源fastjson和Druid。fastjson是JSON解析库，可以解析JSON格式的字符串，支持将Java Bean序列化为JSON字符串，也可以从JSON字符串反序列化到Java Bean。Druid是Java语言中的数据库连接池，能够提供强大的监控和扩展功能，包括监控数据库访问性能、提供WallFilter等。

2012年

- 中国开源云联盟成立。英特尔亚太研发有限公司、新浪网技术（中国）有限公司、中标软件有限公司以及上海交通大学正式签署协议，联合成立“中国开源云联盟”，该联盟是国内首个联合产学研各方组建的开源组织。旨在按照国际上OpenStack开源社区的工作方针，整合中国OpenStack开发者和中国公司的研发资源。
- 腾讯发布首个对外开源项目Tinker。Tinker作为腾讯的第一个开源项目，是Android热解决方案库，支持在不重新安装apk的情况下对dex、library和resources进行更新，推动了Android产品和社区的发展。

2013年

- 三方合作开发Ubuntu Kylin。由中国工业和信息化部下属中国软件与集成电路推广中心（CSIP）、Canonical、国防科技大学三方组成的“CCN开源联合促进实验室”成立，合作开发面向中国用户的Ubuntu版本——Ubuntu Kylin（优麒麟）。
- OSCHINA上线了Gitee（码云）。截至2022年开发者超过800万，托管项目超过2000万，覆盖几乎所有本土原创开源项目。
- CSDN推出代码托管协作平台“CODE”，后升级为GitCode平台，提供全新代码协作管理功能。平台结合CSDN 3500万开发者用户基础及海量知识文档库，通过与GitHub的项目联动，为中文社区开源开发者使用开源、发起开源、推广开源提供专业易用的工具平台及运营服务。

- W3C（万维网联盟）将北京航空航天大学的中国办事处升级为全球总部。2003年，W3C首次在中国举办“中国国际万维网发展论坛”。2011年，百度加入W3C，成为国内首个加入W3C的互联网企业。随后，腾讯、360等互联网企业纷纷加入。

2014年

- 2014年10月16日，由中国支持开源的企业、社区及个人所组成的开源社正式成立，创始成员包括CSDN、GitCafe、GNOME.Asia、微软开放技术（上海）有限公司、优麒麟社区等20余个，之后于2017年改为由个人成员组成。开源社是OSI在中国的首个成员，自2016年起连续举办中国开源年会（COSCon），持续发布《中国开源年度报告》，联合发起了“中国开源先锋榜”、“中国开源码力榜”等，在海内外产生了广泛的影响力。

2015年

- 国务院提出要大力发展开源社区。国务院在《关于积极推进“互联网+”行动的指导意见》中提出大力发展开源社区，鼓励企业自主研发和国家科技计划（专项、基金等）支持形成的软件成果通过互联网向社会开源。引导教育机构、社会团体、企业或个人发起开源项目，积极参加国际开源项目，支持组建开源社区和开源基金会。鼓励企业依托互联网开源模式构建新型生态，促进互联网开源社区与标准规范、知识产权等机构的对接与合作。

2016年

- 国务院支持开源社区创新发展。国务院在《“十三五”国家信息化规划》中提到，引导和支持产学研用深度融合，推动龙头企业和科研机构成立开源技术研发团队，支持科技型中小企业发展，构建产学研用协同创新集群。支持开源社区创新发展，鼓励我国企业积极加入国际重大核心技术的开源组织，从参与者发展为重要贡献者，在优势技术领域争当发起者，积极维护我国相关标准专利在国际开源组织中的权益。
- 云计算开源产业联盟（OSCAR）成立。由工业和信息化部信息化和软件服务业司指导，中国信息通信研究院联合多家云计算开源技术公司发起，中国通信标准化协会代管的云计算开源产业联盟（OSCAR）成立，该联盟是业界首个专注于云计算市场的开源产业联盟。
- 中国首个自主研发的产业级深度学习平台飞桨（PaddlePaddle）正式开源。飞桨以百度多年的

深度学习技术研究和业务应用为基础，是中国首个自主研发、功能完备、开源开放的产业级深度学习平台。

2017年

- 《新一代人工智能发展规划》提出开源开放推进人工智能。7月20号，国务院发布的《新一代人工智能发展规划》提出中国推进人工智能的四条原则：科技引领、系统布局、市场主导、开源开放。其中在开源开放的原则下，共同建设一个体系化的AI技术体系。
- 新一代人工智能产业技术创新战略联盟（AITISA）成立。由北京大学、中关村视听产业技术创新联盟等联合倡议，在科技部试点联盟——数字音视频编解码（AVS）产业技术创新战略联盟的基础上，新一代人工智能产业技术创新战略联盟成立。
- 阿里巴巴捐赠给Apache社区的开源项目RocketMQ正式毕业，成为Apache顶级项目（TLP）。自此，RocketMQ成为国内首个互联网中间件的Apache顶级项目。
- 百度正式对外开源Apollo自动驾驶。Apollo开放平台可以帮助汽车行业及自动驾驶领域的合作伙伴结合车辆和硬件系统，快速搭建一套属于自己的自动驾驶系统。截至2021年底，该平台已完成11个版本迭代，全球拥有135个国家超80000名开发者，合作伙伴超210个，开源代码量总计70万行。
- BCOS开源。BCOS由微众银行、万向区块链、矩阵元三方研发，是聚焦于企业级应用服务的区块链技术平台，于2017年7月31日完全开源。
- Linux基金会授予陆首群“推进开源终身成就奖”。Linux基金会首次在中国召开LC3（LinuxCon&ContainerCon&CloudCon）会议。会上，Linux基金会授予中国开源软件推进联盟名誉主席陆首群“推进开源终身成就奖”。

2018年

- 国务院支持建设一批能够融入国际化发展的开源社区。国务院在《关于深化“互联网+先进制造业”发展工业互联网的指导意见》中提出，面向关键技术和平台需求，支持建设一批能够融入国际化发展的开源社区，提供良好开发环境，共享开源技术、代码和开发工具。
- 国家部委首次以国家科技专项支持开源生态建设。科技部下发国家重点研发计划专项——《云计算与大数据开源社区生态系统》项目，这是国家部委首次以国家科技专项方式支持系统化地推

进开源社区生态系统建设。

- 中国人工智能开源软件发展联盟成立。由工业和信息化部、北京市人民政府共同主办，中国电子技术标准化研究院、中国人工智能开源软件发展联盟（筹）具体承办下，中国人工智能开源软件发展联盟成立大会召开，宗旨是推动我国人工智能开源软件技术和产业发展。
- 中国RISC-V产业联盟成立。由国内外RISC-V领域重点企业、研究机构和行业协会发起的中国RISC-V产业联盟正式成立，旨在解决中国RISC-V领域共同面对的关键问题。包括：建立中国国产自主、可控、安全的RISC-V异构计算平台，促进形成贯穿IP核、芯片、软件、系统、应用等环节的RISC-V产业生态链。
- 中国开放指令生态（RISC-V）联盟成立。由中科院计算所、北京大学、清华大学等近20家国内研究机构和企业联合发起的中国开放指令生态（RISC-V）联盟成立，中国工程院院士倪光南担任理事长。
- 3个原创于中国的项目进入了CNCF。三个项目分别是：Harbor，开源镜像仓库项目，用于容器镜像管理；TiKV，开源分布式事务键值数据库；Dragonfly，由阿里巴巴开源的云原生镜像分发系统。
- 专为物联网设计的轻量级开源虚拟化管理软件ACRN推出。该项目完全由Intel中国开源技术中心实现，从架构设计到团队管理，所有参与项目的工程师都在上海和北京。ACRN为IOT和汽车量身订作虚拟化软件，是该业界第一个开源的软件项目。
- 云原生计算基金会CNCF授予陆首群“开源领袖奖”。CNCF执行董事Dan Kohn代表CNCF在上海召开的首届CNCF&KubeCon国际开源会议上向中国开源软件推进联盟名誉主席陆首群颁发“开源领袖奖”，以表彰他对中国开源事业乃至全球开源事业的突出贡献。

2019年

- OpenCV中国团队成立。引导中国开发力量为OpenCV作贡献，同时改进OpenCV，使之适应未来应用，特别是机器人相关应用，促进科研成果快速应用。
- 木兰开源社区建立。由中国电子技术标准化研究院牵头建立的木兰开源社区，是国家重点研发计划重点专项“云计算和大数据开源社区生态系统”的成果之一，旨在促进产学研用各方开源领域的交流，推动国家科技创新成果开源等。

- **OpenI启智平台成立。**OpenI启智是新一代人工智能产业技术创新战略联盟组织产学研用各界，通力协作、共建共享的开源软件、开源硬件、开放数据社区。
- **头歌社区建立。**该社区主要为高校和企业提供开放的科教项目和课程资源在线开发与服务环境，支持五大类课程/金课建设、实验教学环境建设、科研创新环境建设、工程教育专业认证体系建设等工作。
- **数字孪生体联盟成立。**由工业4.0研究院牵头发起成立数字孪生体联盟，通过运行数字孪生体创新中心，重点建设开源社区和发掘应用场景。
- **SODA Foundation中国区启动。**Linux基金会数据与存储领域的首个子基金会SODA Foundation中国区正式启动，旨在通过智能开放的数据自治标准化建设，帮助用户应对数据基础设施云化转型过程中的挑战。
- **Angel成国内首个从LF AI基金会毕业项目。**腾讯自主研发的Angel从Linux基金会LF AI毕业，成为国内首个从LF AI基金会毕业的项目。
- **华为提出软件开源战略。**华为自研操作系统鸿蒙和方舟编译器开源，宣布未来会将服务器操作系统EulerOS、AI计算框架MindSpore和单机版数据库GaussDB OLTP全部开源。
- **全球Star数最高的React组件库项目。**阿里开源的企业级UI设计语言Ant Design在GitHub的Star数突破5万。2019年7月，Ant Design的GitHub star数超过Material UI，成为全球Star数最高的React组件库项目。
- **阿里开源芯片——无剑100 Open。**平头哥公司推出开源物联MCU芯片平台——无剑100 Open，包含了玄铁处理器、基础接口、开发环境和OS的开源芯片平台。希望能把整个芯片研发周期缩短50%，开发成本降低50%，被阿里定位为面向下一代的AIoT产品平台。
- **全球首个开放云原生应用模型OAM正式开源。**首个开放云原生应用模型OAM（Open Application Model）由阿里巴巴和微软联合开源，这是业界第一个云原生应用标准定义与架构模型。
- **全球首个通用安全计算平台Teaclave进入Apache孵化。**由百度主导研发的Teaclave进入Apache孵化，该平台基于硬件安全能力（Intel SGX），确保敏感数据在可信域外和离岸场景下安全可控的流通和处理。2020年10月，Teaclave社区正式通过并发布了0.1.0版。

2020年

- OSI批准MulanPSL-2.0。2019年，中国首个开源许可证正式发布，名为木兰宽松许可证（Mulan Permissive Software License，Mulan PSL）。2020年，木兰宽松许可证（第2版）正式成为国际化开源许可证。
- Linux基金会宣布成立OpenSSF。OpenSSF可以实现跨行业协作，通过构建具有目标计划和最佳实践的更广泛社区，将领导者聚集在一起，以提高开放源码软件的安全性。
- 开放原子开源基金会成立。由阿里巴巴、百度、华为等十家龙头科技企业联合发起的开放原子开源基金会成立，填补了中国没有开源基金会的空白。
- 上海白玉兰开源开放研究院揭牌。上海白玉兰开源开放研究院是由上海交通大学牵头建设，目标是推动人工智能开源产品的国际规则互认，催生国际开源生态网络关键节点等。
- 清华大学发布自主研发开源时间序列数据管理系统Apache IoTDB。为解决工业互联网领域的时间序列大数据管理需求，清华大学软件学院王建民带领团队自主研发了面向工业领域的开源时间序列数据管理系统Apache IoTDB，成为中国高校首次在Apache国际开源社区发布的软件项目。
- 华为关系型数据库openGauss开源。该数据库在ARM架构芯片上深度优化，并兼容X86架构，采用木兰宽松许可证（MulanPSL-2.0）开源。
- 阿里云开源Kubernetes上第一个云原生平台构建引擎，以及完整实现的OAM模型——KubeVela。KubeVela的开源和规模化落地证明了以OAM为基础的开放应用架构是一条能够充分释放云原生潜力的有效路径，推动OAM成为中国信通院主导发布的《云计算开放应用模型》标准框架。
- 阿里巴巴开源Easy系列算法框架EasyTransfer。阿里云机器学习PAI开源了业界首个面向NLP场景的深度迁移学习框架EasyTransfer。
- 腾讯云企业级分布式KV数据库Tendis开源。Tendis兼容Redis核心数据结构与接口，可提供大容量、低成本、强持久化的数据库能力，适用于兼容Redis协议，需要大容量且较高访问性能的温冷数据存储场景。
- ALC-Beijing向Apache软件基金会申请并成立。ALC-Beijing是面向北京的Apache本地开源社区组织。目前该组织包含10多个Apache顶级项目，包括Kylin、Eagle、RocketMQ、ServiceComb、Griffin等项目。

2021年

- 开源进入国家“十四五规划”。在《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》中，“开源”被首次提及，标志着发展开源成为我国十四五期间的重要工作之一。随后，中共中央、国务院印发的《知识产权强国建设纲要（2021–2035年）》《国家标准化发展纲要》等国家级发展纲要以及《“十四五”国家知识产权保护和运用规划》《“十四五”数字经济发展规划》《“十四五”国家信息化规划》等“十四五”规划，以及工信部、中央网信办、科技部等部委发布印发的《关于加快推动区块链技术应用和产业发展的指导意见》《物联网新型基础设施建设三年行动计划（2021-2023年）》《关于规范金融业开源技术应用与发展的意见》等一系列政策文件，都将开源作为技术创新的主流模式，在多个方面明确提出支持开源发展。
- 工信部将开源纳入重点工作。工信部发布《“十四五”软件和信息技术服务业发展规划》，将“开源重塑软件发展新生态”作为“十四五”期间我国软件产业的四大发展形势之一进行重点阐述。明确提出“建设2-3个有国际影响力的开源社区，培育超过10个优质开源项目”的发展目标，并把“繁荣国内开源生态”作为工作任务，设置“开源生态培育”为专项行动，把“培育重点开源项目、建设优秀开源社区、提升开源治理能力”作为专项行动主要工作，把“加强软件源代码检测和安全漏洞管理能力、提升开源代码、第三方代码使用的安全风险防控能力”作为强化安全保障的重要内容，这是国家级产业规划首次把开源单独作为一个重点工作来部署。
- 五部委发布《关于规范金融业开源技术应用与发展的意见》。中国人民银行办公厅、中央网络安全和信息化委员会办公室秘书局、工业和信息化部办公厅、中国银行保险监督管理委员会办公厅、中国证券监督管理委员会办公厅发布《关于规范金融业开源技术应用与发展的意见》。
- 第一届中国开源教育研讨会（SOSEC）召开。中国第一届开源教育研讨会在四川民族学院举行。来自华东师范大学、四川大学、上海交通大学、东北大学、西北工业大学、中山大学、暨南大学、桂林电子科技大学、四川民族学院、中国科学研究院等高校和研究院（所）的40余名专家、学者汇聚一堂，共同探讨开源教育。
- “科创中国”开源创新联合体成立。该联合体是由中国科协科学技术传播中心、中国电子学会、腾讯、百度、麒麟软件、CSDN等36家单位共同发起成立的开放性、非营利、非法人组织。联合体以打造自主开源生态为宗旨，以建设产学研深度融合的开源创新体系为目标。
- 中国计算机学会CCF开源发展委员会正式成立。该委员会旨在推动探索学术共同体主导的开源

发展新途径，聚焦共同打造开源、开放、中立的产学研协同开源创新服务平台，探索建立CCF开源项目孵化机制，培育原始开源创新项目。

- 华为正式发布面向数字基础设施的开源操作系统欧拉（openEuler）。该操作系统可广泛部署于服务器、云计算、边缘计算、嵌入式等各种形态设备，应用场景覆盖IT、CT和OT，实现统一操作系统支持多设备，应用一次开发覆盖全场景。

- 浪潮研发的NewSQL分布式数据库ZNBase开源，开源后使用的新注册商标名称为KaiwuDB。ZNBase参考谷歌Spanner/F1的设计思想，SQL层使用Go语言开发，基于开源CockroachDB修改，重写其商业代码和开源部分代码，代码修改率76%。存储层使用C++开发，采用多模引擎，涵盖结构化、KV、文件存储、时序存储、图存储等。

- 奥星贝斯开源了OceanBase数据库。基于分布式架构和通用服务器，OceanBase实现了金融级可靠性及数据一致性。2020年5月，OceanBase以7.07亿tpmC的在线事务处理性能创造了TPC-C新的世界纪录。

- 百度开源的ECharts正式从Apache毕业，成为顶级项目。ECharts基于JavaScript的数据可视化图标库，可以生产直观、可交互、定制化的数据可视化图表。

- 第四范式的机器学习数据库OpenMLDB开源。OpenMLDB在半年内实现全球代码托管平台GitHubtopics排名。其中，机器学习数据库、AI数据库方向best-match取得Top1；内存数据库、特征存储方向位列Top3。

- 小米推出新一代Kaldi。由小米集团首席语音科学家，原约翰·霍普金斯大学语言和语音处理中心的DanielPovey主导，正式推出了新一代Kaldi，该项目起源于2009年约翰·霍普金斯大学的夏季研讨会。

- 国内首个违反GPL的案件判决生效。广东省深圳市中级人民法院审理罗盒公司诉风灵公司案的一审判决体现了中国法院对开源软件侵权审理思路的转变。在一审中，法院明确指出GPLv3协议是一种民事法律行为，具有合同性质，可以认定为授权人和用户间订立的著作权协议，属于《合同法》调整的范围。此判例称得上是国内首个明确GPLv3协议法律效力的案例，对国内开源软件侵权行为提供了“有法可依”的背书。

- 全球开源技术峰会GOTC2021召开。该峰会由开放原子开源基金会举办，由上海站与深圳站组

成，覆盖云原生、大数据、人工智能、物联网、区块链、DevOps、开源治理等多个技术领域，为开发者带来全球最新的开源技术。

- 木兰开源社区参加“国家‘十三五’科技创新成就展”。木兰开源社区被国家科技部推荐作为新兴技术板块参加“国家‘十三五’科技创新成就展”，这是我国开源领域首次被国家认可，与蛟龙号、高铁等大国重器一起进入国家顶级科技成就展览。

- 浙江省发布全国首个《开源社区知识产权管理规则指引（试行）》，共十二条，包括管理原则、管理平台、社区组成、约束机制、协同研发、软件管理、专利管理、商标管理、风险管理等方面，提出了合法正当、创新引领、应用先导、发展优先以及数据安全的管理原则。

- Elastic License 2.0（ELv2）虽然满足源码可见，但并不符合OSI的开源定义，采用ELv2的StarRocks因自称开源，而引起业界非议。

- 中国信通院建立可信开源标准体系，推动8个行标立项。

- 中国电子技术标准化研究院推动国家开源标准体系顶层设计，推动相关开源标准立项，分别涵盖术语、元数据、许可证框架、开源项目、开源贡献者、开源治理等方面，首个开源领域国家标准《信息技术开源开源许可证框架》获批立项。

- 2021年6月22日，中科院计算所牵头开发的香山开源高性能RISC-V处理器在首届RISC-V中国峰会发布；7月，香山处理器完成第一版（代号雁栖湖）的28nm流片；10月30日，香山处理器团队负责人包云岗研究员在第六届中国开源年会发表演讲《开源芯片的挑战与机遇》；12月6日，香山通过RISC-V Summit第一次在国际RISC-V社区正式亮相。

- 2021年10月19日，阿里平头哥于2021 云栖大会宣布开源四款玄铁RISC-V系列处理器，成为系列处理器与基础软件的全球首次全栈开源。

- 木兰开源社区吸纳OpenDigger项目进入孵化，推动和完善开源项目、社区衡量指标建设。

- 木兰开源社区的木兰宽松许可证已有多达10万个国内项目采用。

- 2021年11月4日，龙蜥操作系统宣布将捐赠到开放原子开源基金会进行孵化。目前，龙蜥操作系统已在阿里云全面上线，总装机量达百万量级。龙蜥社区（OpenAnolis）已拥有500多家生态企业。其中统信软件、中国移动云等已基于龙蜥操作系统发布商业版本。龙蜥也是国内首个从操作系统层面提供全软件栈国密算法的OS解决方案，并且对内核SM4算法做了深度优化，性能提升近

800%，让中国国密算法从合规走向生产应用。

- 2021年11月，华为openEuler正式捐赠给开放原子开源基金会。
- 2021年10月28日，开放原子开源基金会技术监督委员会投票通过开源项目OpenCloudOS进入开放原子开源基金会孵化。
- 2021年12月9日，由赛昉科技支持的RISC-V开源社区平台RVspace正式上线。
- 2021年12月22日，国产开源操作系统OpenCloudOS的开源社区正式成立。腾讯、宝德、北京初心、北京红旗、飞腾、浪潮、龙芯中科、OPPO、先进开源、中电科申泰、中科方德、兆芯等20余家操作系统生态厂商及用户成为首批创始单位。

2022年

- 2022年2月，TODOGroup与Linux基金会研究合作，发布了《开源项目办公室的演变》白皮书。
- 2022年3月，开源社ONESGroup官宣成立，专注企业开源战略研究。
- 2022年3月，姜宁当选Apache基金会董事、堵俊平再次当选LFAI & DATABoard主席。
- 2022年4月，TiDB宣布正式开源分析引擎TiFlash。TiFlash，基于ClickHouse开发，是TiDB HTAP形态的关键组件，它是TiKV的列存扩展，通过RaftLearner协议异步复制，但提供与TiKV一样的快照隔离支持。这个架构解决了HTAP场景的隔离性以及列存同步的问题。自5.0引入MPP后，也进一步增强了TiDB在实时分析场景下的计算加速能力。
- 2022年4月，国家又进一步提出了《知识产权强国建设纲要和“十四五”规划实施年度推进计划》、《中国银保监会办公厅关于银行业保险业数字化转型的指导意见》，到了12月，三部门联合发布的《互联网信息服务深度合成管理规定》，也将对AIGC以及开源领域产生深远影响。
- 2022年5月，中国信通院牵头成立“可信开源合规计划”并发布《开源合规指南（企业篇）》。《开源合规指南（企业篇）》由中国信通院牵头，联合“可信开源合规计划”的成员单位的专家共同编写完成，分析总结了开源合规发展特点，重点围绕开源软件与数据合规风险、开源许可协议规定、企业开源合规风险防控三大环节展开研究，对于企业做好开源合规具有很强的指导意义，为推动我国开源合规建设提供了参考。

- 2022年5月，阿里巴巴宣布开源的FederatedScope框架，使用事件驱动的编程范式来构建联邦学习，支持大规模、高效率的联邦学习异步训练，能兼容不同设备运行环境，且提供丰富功能模块，降低了隐私保护计算技术开发与部署难度，目前该框架现已面向全球开源。
- 2022年6月，中国首个桌面操作系统开源社区开放麒麟正式发布。开放麒麟（openKylin）由麒麟软件、国家工业信息安全发展研究中心等单位联合成立，目标通过开放操作系统源代码的方式，让更多的开发者共同参与国产开源操作系统的开发，从而打造具有自主创新技术的开源桌面操作系统。
- 2022年6月，石原子宣布开源一体化实时HTAP数据库StoneDB。StoneDB是一款全面兼容MySQL的实时HTAP数据库，自研Tianmu引擎对标MySQLHeatWave架构，具备行列混存、自适应压缩、知识网格和智能索引等核心技术特性，能够高效解决MySQL在分析场景中面临的性能瓶颈，同时支持数据强一致性的事务特性，具备完整的事务并发处理能力，相比MySQL原生分析能力最高可提升100倍。
- 2022年6月，华为开源业界首个分布式云原生开源套件Kurator。分布式云开源套件Kurator助力企业业务跨云跨边、分布式化升级，目标是标准化、统一分布式云管理，整合Karmada、KubeEdge、Volcano、Kubernetes、Istio、Prometheus等业界主流开源技术栈，为用户提供多云、多集群统一编排，统一调度，统一流量治理，边云协同，统一监控运维等核心能力，助力企业业务跨云跨边、分布式化升级。
- 2022年6月，腾讯开源一站式为服务解决方案SpringCloudTencent。SpringCloudTencent实现了SpringCloud标准微服务SPI，开发者可以基于SpringCloudTencent快速开发Spring Cloud云原生分布式应用，实现诸如服务注册与发现、配置中心、服务路由、限流熔断以及元数据链路透传能力等分布式微服务场景。
- 2022年7月，蚂蚁集团宣布开源云原生时序数据库CeresDB。CeresDB是一款高性能、分布式、Schema-less的云原生时序数据库，能够同时处理时序型（time-series）以及分析型（analytics）负载。不同于传统时序数据库，CeresDB的目标不仅仅是能够处理具备常规时序特征（Timeseries）的数据，同时也要能够应对复杂的分析型场景。
- 2022年7月，《2022中国开源发展蓝皮书》正式发布。
- 2022年8月，InfoQ《中国开源发展研究分析 2022》发布。

- 2022年8月，TDengine开源三周年，发布TDengine3.0。TDengine 3.0正式发布，升级成为一款云原生时序数据库（Time Series Database），解决了困扰时序数据库发展的高基数难题，支持10亿个设备采集数据、100个节点，支持存储与计算分离。
- 2022年8月，Tina Tsou当选LFEdgeBoard主席。其中尤为引人瞩目的是：Tina是第一位担任企业互联网工程任务组（IETF）工作组主席的中国女性，也是国际电联电信标准化部门（ITU-T）历史上来自亚洲最年轻的报告人。她还曾担任Akraino技术指导委员会主席。
- 2022年9月，在中国信通院发起的OSCAR开源产业大会上，国内首份《2022年OSPO案例汇编（第一期）》正式发布，收录了多家企业的OSPO实践案例。
- 2022年10月，由Linux Foundation APAC布道者团队OSPO SIG发起的OSPO Summit官宣启动。
- 2022年10月，字节跳动开源自研数据集成引擎BitSail。BitSail支持20多种异构数据源间的数据同步，并提供离线、实时、全量、增量场景下的全域数据集成解决方案。目前，BitSail服务于字节跳动内部几乎所有业务线，包括抖音、今日头条等大家耳熟能详的应用，同时也支撑了火山引擎多个客户的数据集成需求。
- 2022年10月，CSDN《2022中国开源贡献度报告》发布。
- 2022年11月，开放原子开源基金会、X-lab开放实验室和阿里巴巴联合出品的《2022年开源大数据热力报告》发布。
- 2022年11月，格睿云宣布开源云原生时序数据库Greptime。GreptimeDB是分布式、高性能、存储计算分离的开源云原生时序数据库。Greptime使用Rust构建时序数据库。架构方面，Greptime实现了存算分离，以Datanode作为核心组件，具备时序存储引擎、查询引擎和Python Coprocessor等功能，并且功能可以单独划分为特定的计算池，读、写、分析和Python计算的负载相互隔离，互不影响。
- 2022年12月，中国电子技术标准化研究院牵头完成木兰开放作品许可协议和开源治理系列标准。《开源与标准协同发展研究报告》由国内相关专家共同研究和编制完成，从国内外开源项目、社区以及标准研究的实践及成效出发，以更科学的视角系统分析并阐明开源与标准的关系，明确新发展趋势下开源的内涵与外延，为我国后续系统开展开源标准化工作进一步打开思路、给出方向及建议。

2023年

- 2023年01月18日, Apache软件基金会官方宣布由微众银行开源的Apache Linkis顺利毕业, 成为Apache顶级项目(TLP)
- 2023年2月, 由开源中国、华为、南京大学等机构联合发布开源指南针(OSS Compass)平台, 面向GitHub、Gitee等托管平台上托管的所有开源项目, 为国内和国际开源社区提供了SaaS服务, 为中国开源社区的健康指标和结果评估提供量化指标。
- 2023年3月, Linux基金会于中国香港举办了面对面的Linux Foundation Asia Pacific Leadership Summit, 中外专家首次在疫情后面对面的交流。
- 2023年3月3日, Apache软件基金会官网公布了最新一届董事名单, 其中Apache软件基金会孵化器导师、ALC Beijing发起人、华为开源能力专家姜宁成为Apache软件基金会新任董事。
- 2023年4月, Apache软件基金会正式宣布, 由微众银行主导的开源项目Apache EventMesh毕业成为顶级项目(TLP)
- 2023年4月, 由重庆市政府批复设立, 清华大学、华中科技大学、重庆大学、中国信科、中科院等知名高校和企业联合发起成立重庆天工开物开源基金会。
- 2023年4月, InfoQ研究中心发布了《中国开源生态图谱2023》, 共计收录了931个中国开源项目, 涵盖七大细分领域: 操作系统、数据库、人工智能、云原生、大数据、前端、中间件, 以及几大生态机构: 实验室/研究院、开源基金会、开源产业联盟、开发者社区和代码托管平台。
- 2023年5月, 历经半年的激烈角逐, 由中央网信办信息化发展局指导, 中国互联网发展基金会、中国网络空间研究院、中国互联网投资基金联合主办的2022年中国开源创新大赛在北京圆满落幕。此次大赛最终从100多个参选项目中评选出一等奖12个、二等奖26个, 获奖项目覆盖众多技术领域, 部分项目达到国际先进水平。
- 2023年5月26日, 由中国科学院主办的2023年中关村论坛"RISC-V开源处理器芯片生态发展论坛"在北京中关村国家自主创新示范区展示中心举办。本次论坛由重大科技任务局与计算技术研究所联合承办, 北京开源芯片研究院与中国开放指令生态(RISC-V)联盟协办。国内外产学研用各界代表共200余人参会。

- 2023年5月27日，为期2天的全球开源技术峰会在上海张江科学会堂举行，本次峰会由上海浦东软件园、开放原子开源基金会、Linux基金会亚太区和开源中国联合发起，以开源治理、汽车软件、AIGC、AI编程、开源教育培训、云原生、信创等热门话题，探讨开源未来，助力开源发展。
- 2023年6月，LFOSSA举行首个RISC-V人才培养活动，即RISC-V国际人才认证中国区发布会暨数字中国开源之旅-深圳站，发布首个RISC-V认证RVFA的中文版。
- 2023年6月11日，2023开放原子全球开源峰会在北京举行。峰会由开放原子开源基金会、北京市经济和信息化局、北京经济技术开发区管理委员会共同承办，以“开源赋能普惠未来”为主题。
- 2023年 6月15日，第十八届开源中国开源世界高峰论坛会议在北京成功拉开序幕，专家、学者、企业家等开源战略实践者齐聚一堂。COPU主席陆首群致辞，共同探讨开源领域的发展趋势与合作机遇。
- 2023年6月15日，《2023 中国开源发展蓝皮书》正式发布。
- 2023年7月27日，由Linux基金会发起的首次Open Source Congress在瑞士日内瓦成功。开放原子开源基金会出席并参与专题研讨。
- 2023年8月18日-20日，由Apache软件基金会举办的官方全球系列大会CommunityOverCode Asia首次中国线下峰会于北京顺利举行。
- 2023年8月23日，2023 RISC-V中国峰会在北京盛大开幕。本届峰会由北京开源芯片研究院主办，中关村创业大街科技服务有限公司、广州市智能软件产业研究院承办。会议吸引了来自北京大学、清华大学、中国科学院、科技企业、投资机构等2000人参会。
- 2023年8月，RISC-V中国峰会，LFOSSA与RISC-V International首次举办RISC-V中国峰会之国际人才暨开源技术与生态论坛，并联合发布RISC-V人才培训计划，助力培养更多RISC-V在国内的人才。RISC-V International的CEO Calista Redmond在疫情后首次来华并出席RISC-V峰会活动。
- 2023年9月，LFOSSA与CNCNF联合发布云原生人才培养计划。KubeCon + CloudNativeCon + Open Source Summit China于疫情后，首度在上海线下举行。与此同时，LFOSSA与CNCNF联合推出云原生人才培训计划。
- 2023年10月，开源创新教育联盟成立。LFOSSA与中国20多所顶尖的大学、开源社区及机构加入天工开物基金会成立的开源创新教育联盟，一起推动开源教育的创新与发展。

- 2023年10月21日至22日，由中国计算机学会、开放原子开源基金会主办的2023CCF中国开源大会在长沙顺利举行。
- 2023年11月，中国首例AI生成图片著作权侵权案入选2023中国法治实施十大事件。2023年11月27日，北京互联网法院作出一审判决，认为在创作中利用了人工智能技术的涉案图片具备独创性，可以被认定为作品，应受到著作权法保护。双方对该判决均未上诉。判决作出后，多国媒体纷纷跟进报道，判决要旨被译作多种文字。
- 2023年12月举行的Flink Forward Asia大会上，阿里巴巴正式宣布将Flink CDC项目捐赠给Apache基金会，作为Apache Flink的官方子项目。
- 2023年11月15日，经ASF董事会投票决议，时序数据文件格式TsFile正式通过，直接晋升为Apache Top-Level项目（TLP）
- 2023年12月1日，OpenInfra Days China 2023在北京举行。OpenInfra Days China 2023由OpenInfra基金会授权，社区志愿者团队共同筹划组织。
- 2023年12月19日，2023开源产业生态大会在上海召开。本次大会由上海市经济和信息化委员会、上海市科学技术委员会、上海市科学技术协会指导，上海对外经贸大学、上海开源信息技术协会主办，OpenTEKr、DaoCloud道客承办。
- 2023年12月28日，OSCHINA和Gitee联合发布了《2023中国开源开发者报告》，深入剖析了开源领域的技术动态和趋势。

2024年

- 2024年2月8日，开源社正式发布《2023中国开源年度报告》。
- 2024年3月15日，2024年第一期|CCF ODC《开源战略动态月报》
- 2024年3月24日，由开源中国、RISC-V国际人才培养认证中心、厦门市开源芯片产业促进会共同主办，RISC-V国际人才培养认证中心、广东开源芯片产业科技有限公司承办，RISC-V国际基金会、LinuxFoundation开源软件学园、中国开放指令生态(RISC-V)联盟支持的“GDC 2024-RISC-V产业人才培养峰会”在上海举办！

附录 2 2023–2024 年度开源榜单汇编

2023 “科创中国” 开源创新榜单

主办：中国科协科学技术传播中心 中国计算机学会 中国通信学会 中国科学院软件研究所

承办：CSDN

优秀开源项目（20个）

Apollo开放平台、OpenAtome OpenHarmony、SShardingSphere、Wechaty、Nacos、TDengine、jeecgBoot低代码开发平台、Sentinel、Excelize、EMQX、DoKit、长安链.ChainMaker、RT-Thread、deepin（深度）操作系统、OceanBase4.0单机分布式一体化开源数据库、LinkWechat、Alluxio、新一代云原生数据仓库Databend、铜锁开源密码学算法库、Volcano

优秀开源社区（10家）

龙晰社区、openKylin（开放麒麟）社区、openEuler社区、木兰开源社区、Apache DolphinScheduler社区、Apache Doris开源社区、OpenCloudOS下一代云原生开源操作系统、开源无线网络社区、FATE（Federated AI Technology Enabler）开源社区、昇思MindSpore

优秀开源人物（10位）

边思康 马涛 张磊 余跃 代立冬 崔秋 张亮 王泽锋 顾荣 高原

2023年活力开源项目和开源贡献者榜单

主办：开放原子开源基金会

2023活力开源项目榜单

2023生态开源项目

飞桨、龙蜥操作系统/Anolis OS、昇思MindSpore、Apache DolphinScheduler、Apache Doris、Apache Dubbo、Apache EventMesh、Apache IoTDB、Apache Kyubi、Apache ShardingSphere、Apache ShenYu、deepin、EMQX、JeecgBoot、Nacos、OpenAtom openEuler、openGauss、OpenHarmony、openKylin（开放麒麟）、StarRocks、TDEngine、TiDB

2023快速成长开源项目

魔搭FaceChain、新一代Kaldi、隐语SecretFlow、Apache SeaTunnel、Apache StreamPark、Bytebase、Casbin、CeresDB、Curve、DDE桌面组件、Eosphoros-ai、GreatSQL、HertzBeat、Hyperf、JuiceFS、Kuasr、KWOK、LinkWeChat、MyBatisPlus、NutUI、openGemini、OpenSumi、OpenTiny、Postcat、Sa-Token、TuGraph、UBML、VisActor、WeChaty

2023年活力开源贡献者榜单

技术、生态贡献

曾令卡 韩光宇 郑志超 罗杨 徐建国 何森森 李梦柳 骆涛 张留杰 颜鑫 要海峰 韩小鹏
张世明 陈圣杰 李瑾 王勤龙 蔡威 李颖 金子威 李明康 陈锐彪 陈梓立 贺张俭 高原
涂旭辉 张亮 郭旭东 刘宇博 侯昊男 连志安 王华杰

技术贡献

肖宇 李伟刚 王备 徐晓健 李鑫 张奥琳 骆融臻 梁志远 张原嘉 贾佳豪 田川 张一乔
刘远乐 郑必城 马占伟 赵军霞 王石 祁金全 杨显钊 李云龙 王维真 任雪冰 王萌 屈博
满天星 江莎 边广泽 章伟星 汤文辉 崔子航 杨翊 卢雨旼 王士忠 詹荣瑞 赵鑫 何凤恩
于敏杰 张润宇 黄子豪 张守忠 陆琦 张永帝 张国荣 董伟

生态贡献

齐宝玮 张旭芳 查浩浩 刘太俊 范劲雨 李根明 李蒙 陈美汝 朱伟 巴延兴 刘晓燕 梁开祝
丽娜 孙悦 程胜春 刘睿 赵易 李志宇 杨志超 付庆 刘克超 刘锐鑫 蒋卫峰 李宝 孙鑫
王潮

2022中国开源创新大赛

主办：中国互联网发展基金会、中国网络空间研究院、中国互联网投资基金

2022优秀开源项目榜单

一等奖

Apache IoTDB、Apache ShardingSphere、Excelize、Fluid、Hyperf、Karmada、OneOS物联网操作系统、OpenEuler发行版（on riscv）、OpenGauss、RustDesk远程桌面、思否开发者社区、序列大数据的智能计算

二等奖

Apache DolphinScheduler、Apache Doris、Apache Kyuubi、Apache StreamPark、Byzer-lang、CloudWeGo、CodePecker 源代码缺陷分析系统-补阙、Curve、FISCO BCOS国产开源联盟生态圈、Genmai、GitCode开源平台、LinkWeChat、Nacos、OpenCloudOS、OpenFunction函数计算平台、Openkylin、Scheme-langserver、TDengine时序数据库、TuGraph图数据库、UKUI桌面环境、Volcano、建木、“金融业生僻字处理”开源项目、龙蜥操作系统、龙蜥社区、深度桌面操作系统

优秀奖

OpenKylin RISC-V开源版本、定时调度器、CodePecker软件成分分析系统（SCA）-“析微”、OpenAPI Log Cat - 基于OpenAPI的Web自动防护工具、SBOM服务、基于开源鸿蒙打造的行业全场景操作系统 -在鸿OS、迪力匡正软件供应链风险分析工具、JEEWMS、Rvcc、Swow、Eulaceura、ChunJun、AI元宇宙虚拟数字人、HummerRisk云原生安全平台、Slime、事件网格、稻香AI、StoneDB：一体化实时开源HTAP数据库、IoTSharp、Apollo、开源问答社区软件 Answer、Hypercrx、凹语言、基于物联网的智能高铁废液循环系统、OpenCV、Wechaty、开源可信3.0基础软件、深度开源社区、Dromara开源社区

开源贡献人物

安旭 陈渝 董国伟 冯铃 郭炜 郭御风 郭悦 胡春明 胡达川 荆琦 李明宇 李欣博 李中华

梁大功 梁志辉 凌晓峰 刘朝 刘宇 马艳军 孟迎霞 宋可为 孙海龙 谭中意 汤炜伟 王庆
杨军 翟艳芬 张梦茹 庄表伟

2023中国开源先锋33人之心尖上的开源人物

主办: SegmentFault 思否、开源社

2023开源人物

陈渝 丁皓(漩涡) 古思为 郭旭东 季敏 李梦 李钰 林旅强 刘军 刘敏 陆首群 马红伟
马如悦 鸟哥(Linux) 聂励峰 祁宁 三咲智子 王荷舒 王建民 王楠 王庆 王小虎 王旭
王晔晖 王泽锋 姚琴 郁志强 张铎 张国锋 张俊霞 章文嵩 张晓辉 朱其罡

附录 3 开源专有名词

开源

开源是促进信息技术创新的重要途径，是将源代码、设计文档或其他创作内容开放共享的一种技术开发和发行模式。技术领域的开源包含源代码+软硬件设计文档、源数据等技术和资源的开放、共享，源代码仍然是开源的主要内容。在开源模式下，通过许可证的方式，使用者在遵守许可限制的前提下，可自由获取源代码等，并可使用、复制、修改和再发布。

自由软件

自由软件（Free Software），是指可以付费或不付费得到GNU软件，“Free”关乎自由，而不是价格。一旦得到了软件，便拥有了使用它的四项特定自由：可以自由按照自己的意愿运行该软件；可以自由复制软件并将其发送给朋友和同事；可以自由通过对源代码的完全控制而改进程序；可以自由发布改进的版本从而帮助社区建设。

商业软件

商业软件（Business Software）是指通过贸易方式向社会公众发行的各种商品化软件。对于商业软件，供应商不仅应该向用户提供程序和使用说明，而且应该向用户提供包括版本更新在内的技术服务，商业软件在知识产权方面的最重要体现是版权（著作权），即商业软件的权利人享有软件及其复制品的发行权。目前国际上通行的商业软件发行方式是向用户提供使用许可。按照这种方式，用户在支付一定费用后，得到的仍不是这份软件复制品的全部权利，而是一种附条件、有限制的权利；即在遵守使用许可协议的条件下，行使由该软件的使用许可协议所规定的若干项权利，如使用、临时复制等。

开源供应链

开源软件供应链是一个实际业务系统，在开发和运行过程中，涉及所有开源软件上游社区（Upstream）、源码包（Source Package）、二进制包（Binary）、包管理器（Package Manager）、存储仓库（Repository），以及开发者（Developer）和维护者（Maintainer）、社区（Community）、基金会（Foundation）等，按照依赖、组合、托管、指导等关系形成的供应链网络。

开源基金会

开源基金会是开源生态中的重要部分。对于非营利性且处于中立位置的开源基金会来说，它们拥有开源项目的知识产权，而且没有任何商业利益，这为贡献者、开发者以及用户提供了良好的协作平台；基金会也会通过技术服务和项目孵化来帮助开发者和开源企业，如软件仓库、问题跟踪、技术指导、法律支持、项目投资、公共关系维护等；基金会还会提供项目日常的运营和治理，以满足在项目生命周期不同阶段对于管理项目的需求。

Apache基金会

Apache软件基金会（Apache Software Foundation, ASF），是专门支持开源软件项目的非盈利性组织。在它所支持的Apache项目与子项目中，所发行的软件产品都遵循Apache许可证（Apache License）。Apache基金会倡导的是合作和贡献，信奉“贡献者→提交者→成员→导师”路径。想成为贡献者，需积极为Apache社区贡献代码、补丁或文档。想成为提交者，需要成员的指定，而成为了提交者，就会拥有一些“特权”，提交者中的积极分子和优秀生可“毕业”成为成员。

Linux基金会

Linux基金会通过提供财务和智力资源、基础设施、服务、活动以及培训来支持创建可持续开源生态系统。Linux是大多数开源软件的基础，该组织的核心目标是推动Linux系统的发展。近年来，随着开源的兴起，Linux基金会适应时代发展的需要，在Linux的基础上扩大涉足领域，并监管大型协作项目。业界Xen、KVM、CNCF、Hyperledger等知名项目，都来自Linux基金会。

开放原子开源基金会

开放原子开源基金会是我国首个开源基金会，主要提供基础服务、法律服务、IT基础设施服务、社区运营管理等四大类别服务。该基金会是由民政部登记、工业和信息化部指导的基金会。开放原子开源基金会设理事会、技术监督委员会与秘书处。理事会负责审议和修改基金会章程等；技术监督委员会是其中立的技术决策机构，负责基金会技术相关的决策及项目的孵化评审等；秘书处是基金会的执行机构，负责基金会日常事务等工作。

云原生计算基金会（CNCF）

云原生计算基金会（Cloud Native Computing Foundation, CNCF）是Linux基金会旗下的基金会，成立于2015年12月11日，其口号是坚持和整合开源技术来让编排容器作为微服务架构的一部分，是致力于云原生应用推广和普及的一支重要力量。CNCF作为一个厂商中立的基金会，致力于GitHub上的快速成长的开源技术如Kubernetes、Prometheus、Envoy等的推广，帮助开发人员更快、更好地构建出色的产品。

开源基础设施基金会（OIF）

2020年10月，OpenStack基金会（OSF）正式演进为开源基础设施基金会（Open Infrastructure Foundation, OIF）。OIF的目标是在全球范围内服务开发者、用户及生态系统，提供共享资源，以扩大OpenStack公有云与私有云的成长，从而帮助技术厂商选择平台，助力开发者开发出行业最佳的云软件。OIF分为个人会员和企业会员两类，个人会员是免费无门槛的，他们可凭借技术贡献或社区建设加入OpenStack社区。企业会员则根据赞助会费的情况，分为白金会员、黄金会员、企业赞助会员及支持组织者，其中白金会员的话语权最大。

OIN（Open Invention Network）专利保护社区

OIN是目前为止最大的专利保护社区，旨在维护关键开源软件（OSS）项目参与者的专利自由，支持开源软件（OSS）关键元素Linux的自由开发环境。核心技术专利互不侵犯符合开源软件内在的文化

常态，通过加入OIN社区，足以了解社区里的行为诚信与否，帮助社区成员降低专利风险。OIN成立时获得了Google、IBM、NEC、飞利浦、索尼、SUSE和丰田等企业的大力支持，拥有3,400多位社区成员和260余万项专利与申请。

开源社区

开源社区是众多创作者和使用者交流和学习的“基地”。开源社区是项目开发的组织形式，是由所有参与开发和改进源代码项目的用户组成的社群，通常也是遵循某项开源许可协议发布源代码的网络平台。在开源社区，众多开发者进行合作与开发，并分享成果。

木兰开源社区

木兰开源社区是国家重点研发计划重点专项“云计算和大数据开源社区生态系统”的成果之一，旨在促进产学研用各方开源领域的交流，推动国家科技创新成果开源，加强企业、科教单位和行业用户之间的沟通，推动开源成果转化落地。该社区为各类开源项目提供中立托管，保证开源项目的持续发展不受第三方影响，通过更加开放的方式来打造和完善开源社区生态。2019年8月，木兰系列开源许可证的《木兰宽松许可证，第1版》正式上线，MulanPSL v2于2020年通过OSI认证，成为首个我国主导的中英双语开源许可证。

CSDN社区

中国开发者社区CSDN (Chinese Software Developer Network)，创立于1999年。合作企业上千家，旗下拥有全球最大的中文开发者社区 CSDN.NET、专业技术博客CSDN BLOG、在线学习平台程序员学院、中国特色的代码托管平台 GitCode、高校合作平台CSDN教育、IT技术人才猎头服务科锐福克斯等，全力为中国IT技术人员提供知识传播、在线学习、职业发展、工具赋能等全生命周期服务，是技术创新、行业创新、中国自主IT生态重塑、开发者最优云服务平台。

开源社

开源社（“KAIYUANSHE”）成立于 2014 年，是由志愿贡献于开源事业的个人成员，依“贡献、共识、共治”原则所组成，始终维持“厂商中立、公益、非营利”的理念，以“立足中国、贡献全球，推动开源成为新时代的生活方式”为愿景，以“开源治理、国际接轨、社区发展、项目孵化”为使命，旨在共创健康可持续发展的开源生态体系。开源社积极与支持开源的社区、高校、企业以及政府相关单位紧密合作，同时也是全球开源许可证认证组织 OSI 在中国的首个成员。自2016年起连续举办中国开源年会（COSCon），持续发布《中国开源年度报告》，联合发起了“中国开源先锋榜”、“中国开源码力榜”等，在海内外产生了广泛的影响力。

开源许可证

开源许可证是对开源技术进行规范的授权合同，也可称为授权协议书，具有合同和著作权的双重法律性质。开源许可证采用契约和授权方式，指导和规范许可人和被许可人在处理开源作品时的权利、义务和责任，是解决开源面临的法律和商业问题的核心机制。

GPL许可证

GPL许可证（GNU General Public License）是自由软件开源联盟GNU开源许可证的一种，同时也是开源软件领域对被许可人权利限制最严的许可证。GPL许可证最大的特点在于它要求根据GPL许可证发布的软件（如Linux操作系统）修改、翻译的演绎作品，甚至只要其中任何一部分代码是以GPL发布的，那么全部程序也必须受GPL许可证的约束，即继续遵守GPL许可证的规定。

BSD许可证

相对于其他开源软件许可证，特别是GPL许可证，BSD（Berkeley Software Distribution）许可证对被许可人来说可能是最“宽容”的。虽然BSD许可证具备开源软件许可证普遍的要求，但BSD许可证只要求被许可者附上该许可证的原文以及所有开发者的版权资料。也就是说，只要标明了源代码的出处，被许可人可以将其用在自己的软件中，并按自己的要求（包括以商业软件的方式）再发布或

再许可等。因此，BSD许可证在学校或公共科研机构研发的开源软件转化为产品方面发挥了重要作用。

LGPL许可证

LGPL许可证 (Lesser General Public License或Library General Public License) 中文可译为“较宽松公共许可证”或“函数库公共许可证”，也是自由软件联盟GNU开源软件许可证的一种。在具体执行 LGPL许可证时，与GPL许可证最大的不同之处在于其适用于特殊设计的函数库，并允许非自由的程序可以与这些函数库连接。

MPL许可证

MPL (Mozilla Public License) 最初是1998年Netscape的Mozilla小组为其开放源代码软件项目设计的软件许可证。与GPL许可证和BSD许可证相比，MPL的许多权利与义务与它们相同，但也存在不同之处，比如，许可证允许被许可人将经过许可证获得的源代码同自己其他类型的代码混合得到自己的软件程序。

最终用户许可协议

最终用户许可协议 (End User License Agreement, EULA) 指的是一家公司的软件与软件的使用者所达成的协议，是软件应用程序作者或者发布者与应用程序使用者之间的合法合同。最终用户许可协议对软件的使用、修改和分享及其他相关事宜作出规定。

开源代码托管平台

开源代码托管平台，简单来说就是存储、管理、维护源代码，促进项目协同开发的网络平台。现在代码托管技术以Git为主，大多数企业使用Git进行协同开发，很多代码托管平台也是通过Git来实现的，如GitHub、GitLab、Gitee、GitCode、GitLink、AtomGit。

GitHub

GitHub致力于支持世界各地的软件创新，是所有开发者的家园。软件质量和软件供应链的安全取决于所有人的贡献，GitHub服务全球的一致性和可用性是重中之重。参与全球软件协作有益于合作创新，并支持政府在数字主权、安全和可持续性方面的目标。

技术

世界知识产权组织在1977年版的《供发展中国家使用的许可证贸易手册》中将技术定义为：“技术是制造一种产品的系统知识，所采用的一种工艺或提供的一项服务，不论这种知识是否反映在一项发明、一项外观设计、一项实用新型或者一种植物新品种，或者反映在技术情报或技能中，或者反映在专家为设计、安装、开办或维修一个工厂或为管理一个工商业企业或其活动而提供的服务或协助等方面。”该定义是至今为止国际范围内对“技术”一词最全面、完整的解释。知识产权组织把世界上所有能带来经济效益的科学知识都定义为技术。

RISC-V

开源指令集架构是一个基于精简指令集（RISC）原则的开源指令集架构（ISA）。与大多数指令集相比，RISC-V指令集可以自由地用于任何目的，允许任何人设计、制造和销售RISC-V芯片和软件。适用于现代计算设备（如仓库规模云计算、高端移动电话和微小嵌入式系统）。

开放科学

根据联合国教科文组织（UNESCO）在2021年11月发布的《开放科学建议书》，开放科学被定义为一种包容性的结构，它结合了各种运动和实践，旨在使多语言科学知识对所有人开放、可访问和可重复使用，以增加科学合作和信息共享，以造福于科学与社会，并向传统科学界之外的社会行为者开放科学知识创造、评估和交流的过程。开放科学的思想最早萌芽于中世纪的欧洲宫廷资助系统，资助者采用同行评议、开放交流的方式鉴别受资助成果的质量。《中华人民共和国科学技术进步法》（2021年修订）第九十五条规定“推动开放科学的发展，促进科学技术交流和传播”。

开放标准

标准记录了产品或相关工艺和生产方法的规则、条件、准则或特性的使用。标准可以来自一家公司，其成功的产品成为“事实上的”标准。标准也可以由两个或两个以上的软件生产商之间的协议制定，或者通过政府法令或政府运行的流程，或者由多个政府的代表制定。在政府流程中制定的标准，例如国际标准化组织（ISO）的标准，称为“法律”标准。或者，可以通过成员的包容性共识过程来制定标准，并按照明确定义的政策和程序进行排序。这些标准由协商一致进程中的公共和私营部门参与者商定，他们通常同意以非歧视性的方式免费提供其标准。这些标准被称为“开放”标准。

附录 4 开放科学和开放标准

开放科学

开放科学的发展历程开放科学的思想最早萌芽于中世纪的欧洲宫廷资助系统，资助者采用同行评议、开放交流的方式鉴别受资助成果的质量。开放科学的发展大致经历了三个阶段。欧洲一直走在开放科学的前沿，2009年欧盟第七框架计划（FP7）资助了欧洲开放获取基础设施项目（OpenAIRE），致力于建设开放和可持续的学术交流基础设施。2014年，在欧洲“地平线2020计划”下提出了欧洲开放科学培训项目（FOSTER）。2018年，欧盟支持发起开放获取S计划，旨在使得受公共和私人支持的，经过同行评议的出版物能够完全和立即开放获取。美国开放科学中心（COS）于2013年发布了开放科学框架（OSF），旨在协助科研团队管理项目和公开成果。《中华人民共和国科学技术进步法》（2021年修订）第九十五条规定“推动开放科学的发展，促进科学技术交流和传播”。

根据联合国教科文组织（UNESCO）在2021年11月发布的《开放科学建议书》，开放科学被定义为一种包容性的结构。开放科学包括所有科学学科和学术实践的各个方面，包括基础科学和应用科学、自然科学和社会科学以及人文科学，它建立在开放科学知识、开放科学基础设施、科学传播、社会行为者的开放参与并与其他知识系统进行公开对话的框架之上。

开放科学的核心价值观、指导原则和行动领域，《开放科学建议书》概述了开放科学的四大核心价值观、六项指导原则、七个行动领域。

开放科学的四大核心价值

- （1）质量和完整性：开放科学应该尊重学术自由和人权，支持高质量的研究，方法是汇集多种知识来源，广泛提供研究方法和成果，供严格审查和透明的评估过程使用。
- （2）集体利益：科学知识应该是公开且普遍共享的。科学实践应该是包容的、可持续的和公平的，也应该是科学教育和能力发展的机会。
- （3）公平公正：开放科学应发挥重要作用，确保发达国家和发展中国家科学家之间的公平，实现科学投入和产出的公平互惠共享，使知识的生产者和消费者都能平等获得科学知识。
- （4）多样化与包容性：开放科学应该包含知识、实践、工作流程、语言、研究成果和研究主题的多样

性，以支持科学界作为一个整体的需求和知识多元化，支持多样化的研究团体和学者，支持传统科学界之外的更广泛的公众和知识持有者，包括土著人民和当地社区，以及不同国家和地区的社会行动者。

开放科学的六项指导原则

- (1) 透明度、审查、批评和可重复性；
- (2) 机会均等；
- (3) 责任、尊重和问责；
- (4) 合作、参与和包容；
- (5) 灵活性；
- (6) 持续性。

开放科学的七个行动领域

为实现开放科学建议书的目标，建议成员国根据国际法并考虑到各自的政治、行政和法律框架，在以下七个领域同时采取行动。

- (1) 促进对开放科学、相关利益和挑战以及开放科学的多种途径的共同理解；
- (2) 为开放科学营造有利的政策环境；
- (3) 投资开放科学基础设施和服务；
- (4) 投资于开放科学的人力资源、培训、教育、数字素养和能力建设；
- (5) 培育开放科学文化，调整开放科学激励机制；
- (6) 在科学进程的不同阶段促进开放科学的创新方法；
- (7) 在开放科学的背景下促进国际和多方利益攸关方合作，以减少数字、技术和知识差距。

开放获取

2002年,《布达佩斯开放获取倡议》(BOAI)正式提出开放获取概念。2003年,德国马普学会等机构发起开放获取柏林会议,并发表《柏林宣言》,呼吁免费公开更多的科学资源,鼓励科学家以OA方式发表论文。2004年, Springer出版集团开始在其1000多种订阅期刊上向作者提供开放获取选项,带动期刊向符合开放获取期刊转换。2006年,高能物理开放获取联盟SCOAP 3建立,推动学科内大多数期刊向开放获取出版转换。2007年,南非开普敦开放教育会议发布了《开普敦开放教育宣言》。2009年,开放获取知识库联盟COAR成立,旨在联合遍布五大洲的机构知识库并建立一个全球协作网络,提高存储成果的可见度和利用度。2013年,全球研究理事会发布《出版物开放获取行动计划》。2016年3月 德国马普学会联合多家机构发起“OA 2020计划”。2018年9月4日,在欧盟委员会和欧洲研究理事会(ERC)支持下成立S联盟(Coalition S),发布《加速向全面且立即的科研出版物开放获取转变倡议》。

开放获取的主要方式

《布达佩斯开放获取倡议》提出了实现开放获取的两条路径,即开放获取期刊(Open-access Journals)和开放存储(Self-archiving),这两种方式分别被称为金色OA和绿色OA。除这两种主要方式外,在实际出版过程中还有其他的开放获取模式。

- 金色OA (GoldOA): 指论文的最终版本在网络上立即开放获取,读者不需要付费。金色OA的论文一般发表在完全开放获取或向开放获取转换的期刊上,遵循CCBY或CCBY-NC-ND的版权许可协议,由作者、研究机构或基金资助机构缴纳论文处理费(Article Processing Charge,简称APC)。
- 绿色OA (GreenOA): 指作者在订购期刊上发表论文后,将论文一定版本存到开放获取知识库中并在一定时间后(Embargo Period,开放时滞期)开放获取。
- 青铜OA (BronzeOA): 指科研论文的许可并不明确,出版商可能限时授予文献免费访问权限的获取方式。
- 钻石OA (DiamondOA): 指在不收取作者费用的情况下,发布开放获取的期刊,即作者和读者双向免费。2022年3月2日, Science Europe、S联盟(cOAlition S)、人文学科开放学术交流研究基础设施服务联盟项目(OPERAS)和法国国家研究机构(ANR)共同提出了“钻石OA行动计划”。
- 黑色OA (BlackOA): 是较新的说法,指的是从学术社交网(如Mendeley、ResearchGate、

Academia.edu等)或非法提供学术论文全文的网站(如Sci-Hub)上可以免费下载大量本应交费浏览的学术文章。

开放获取的载体

CORE (<https://core.ac.uk/>)是全球最大的开放获取研究论文库,其使命是索引全球所有开放获取研究,并为所有人提供不受限制的访问。

Dimensions是最全面的研究资助数据库,它将资助与数百万种由此产生的出版物、临床试验和专利联系起来。

Web of Science (<https://www.webofscience.com/>)是全球最受信赖且独立于出版方的全球引文数据库。

开放获取期刊指用户可以在法律允许的范围内最大限度地接近学术资源,打破传统学术期刊的限制和门槛,排除经济、地域等因素,按照个体需求进行阅读和分享学术期刊。美国科研出版社(Scientific Research Publishing,简称SCIRP)是全球最大的开放获取期刊出版商之一。目前已出版超过200多种开放获取、在线、同行评审的期刊,涵盖了广泛的学术学科。此外,国际上开放出版期刊平台还有PLOS、Cogent OA、Dove Medical Press、Copernicus Publications、ELCVIA、Frontiers等。DOAJ是全球开放获取期刊索引,只接受开放获取期刊,期刊应遵守学术出版的透明度和最佳实践原则,期刊必须有自己的专用URL和主页,必须至少有一个ISSN,期刊必须有编辑和编委会,所有文章在发表前必须通过同行评审。

机构知识库(Institutional Repository,简称IR)又称机构典藏库、机构仓储、机构库等,通常情况下指的是,某个机构实体为保存和管理本机构智力成果资源(包括本机构成员创作的学术成果、科研数据、课件教程、软件工具等),自主或借助第三方机构建立的集提交、组织、检索、分析、利用等为一体的学术存储与服务系统。OpenDOAR是开放获取存储库全球目录,截至2024年5月,OpenDOAR平台注册的机构知识库总数达5,907个,其中美国(818个)、日本(655个)、德国(305个)排名前三位,中国大陆66个,占比1.11%。国际上机构知识库联盟主要包括:开放获取知识库联盟(COAR)、法国FAL、荷兰DAREnet、澳大利亚ARROW、日本JAIRO、德国OA-Network、英国JISC RepositoryNet、欧盟DRIVER。我国的机构知识库联盟主要有中国科学院机构知识库网络(CASIRGRID)、中国高校机构知识库联盟(Confederation of China Academic Institutional

Repository, 简称CHAIR) 和农业机构知识库联盟。

预印本 (Preprint) 是作者在提交期刊出版前未经严格同行评议的手稿 (投稿版、录用稿), 经过初步的评议审核后, 即借助于预印本平台在最短的时间内以开放获取的形式发布, 为作者获得最新研究成果的网络首发权, 为学术交流系统提供最新、最快速的研究成果传播与利用渠道。学术论文出版的预印本平台有: arXiv、BioRxiv、medRxiv、ChemRxiv、F1000、figshare、PeerageofScience、engrXiv、PeerJPreprints等。其中, arXiv于1991年发起, 涉及物理学、数学、计算机科学、定量生物学、定量金融学、统计学、电气工程和系统科学以及经济学领域的超过200万篇学术论文。BioXiv、MedXiv和ChemXiv分别是生命科学、医学和化学领域的预印本平台。ChinaXiv是中国科学院文献情报中心于2016年开始运营的预印本交流平台, 支持中英文科技论文预印本的发布、传播、下载和评论。

开放标准

标准记录了产品或相关工艺和生产方法的规则、条件、准则或特性的使用。标准可以来自一家公司, 其成功的产品成为“事实上的”标准。标准也可以由两个或两个以上的软件生产商之间的协议制定, 或者通过政府法令或政府运行的流程, 或者由多个政府的代表制定。在政府流程中制定的标准, 例如国际标准化组织 (ISO) 的标准, 称为“法律”标准。或者, 可以通过成员的包容性共识过程来制定标准, 并按照明确定义的政策和程序进行排序。这些标准由协商一致进程中的公共和私营部门参与者商定, 他们通常同意以非歧视性的方式免费提供其标准。这些标准被称为“开放”标准。

开源促进会 (OSI) 是负责审查和批准符合开源定义 (OSD) 许可证的组织, 该组织表示: “开放标准”不得禁止在开源软件中实现符合要求。OSI 列出了开放标准必须满足的五个基本要求 (<https://opensource.org/osr>):

(1) 没有故意的秘密。标准不得隐瞒可互操作实现所需的任何细节。由于缺陷是不可避免的, 因此该标准必须定义一个过程, 用于修复在实施和互操作性测试中发现的缺陷, 并将上述更改纳入该标准的修订版本或替代版本, 以在不违反OSR的条款下发布。

(2) 可用性。该标准必须以合理和非歧视性的成本在免版税条款下免费和公开提供 (例如, 从稳定的网站)。

(3) 专利。所有对实施标准至关重要的专利必须: 根据免版税条款获得许可, 不受限制地使用; 或在

开源软件开发时予以清晰明确的使用声明。

(4) 无协议。不得要求执行许可协议、NDA、授权、通过点击或任何其他形式的文书工作来部署符合标准的实施。

(5) 没有与开放标准要求不兼容的依赖关系。标准的实施不得要求任何其他不符合此要求标准的技术。

开放标准的原则

“开放标准”原则构成了全球市场经济的现代化模式，以技术创新为动力源泉并推动了标准的全球化部署，而无论这些标准在传统的国内体系中处于何种地位。该声明的5个核心原则要求如下：

- (1) 标准机构间相互协作；
- (2) 在标准制定中，坚持遵守流程、达成广泛一致、公开透明和利益平衡；
- (3) 致力于推动技术先进性、互操作性、相互竞争、创新性和为人类谋福利；
- (4) 标准规范的可得可用；
- (5) 自主决定是否采用。

在政府治理方面，英国政府针对对软件、数据和文档格式互操作的要求，提出了开放标准的7条原则：

- (1) 开放标准必须满足用户需求；
- (2) 开放标准必须使供应商能够平等地获得政府合同；
- (3) 开放标准必须支持灵活性和变化；
- (4) 开放标准必须支持可持续成本；
- (5) 使用明智的决策选择开放标准；
- (6) 使用公平透明的流程选择开放标准；
- (7) 使用公平透明的流程指定和实施开放标准。

常见的开放标准组织

- 标准机构: DMTF (分布式管理任务组公司)、IETF (互联网工程任务组)、NIST (美国国家标准与技术研究院)、OASIS (结构化信息标准促进组织) 和The Open Group (开放组小组)
- 软件标准: C++ 标准委员会、DWARF 调试标准、Java 社区进程、OSGi联盟、UEFI (统一可扩展固件接口论坛)、OpenMP、PatternFly
- 安全标准: CSA (云安全联盟)、Trusted Computing Group (可信计算组)
- 基准标准: SPEC (标准绩效评估公司)、STAC (安全技术分析中心)、TPC (事务处理绩效委员会)

开放科学数据

开放科学数据包括具有以下特点的原始的和经过处理的数字数据与模拟数据及附属的元数据, 以及数字分数、文字记录、图像和声音、协议、分析代码和工作流程等: 任何人都可以公开使用、重用、保留和重新发行, 但须注明出处; 依照良好数据治理和管理原则, 特别是FAIR (可查询、可获取、可互操作和可重复使用) 原则, 及时以方便用户、人机可读和可操作的格式提供, 并辅以定期管理和维护。

FAIR原则: 2014年1月, 一些代表学术界、工业界、资助机构和学术出版商的多元化利益相关者共同研讨起草了一份指导科学数据管理的倡议性文件, 倡导科研活动产出的数据在开放共享过程中应该努力实现可发现 (Findable)、可访问 (Accessible)、可互操作 (Interoperable) 和可重用 (Reusable), 4个目标层层递进, 简称为“FAIR原则”。2016年, FORCE11 (一个由学者、图书馆员、档案管理员、出版商和研究资助者组成的社区, 将4个抽象目标细化为15条具体原则正式发布。

CARE原则: 土著数据治理原则 (The CARE Principles for Indigenous Data Governance, 简称CARE) 四个字母分别对应“关注集体利益 (Collective Benefit)”、“控制权 (Authority to Control)”、“责任 (Responsibility)”和“伦理 (Ethics)”, 由全球土著数据联盟 (GIDA) 于 2019 年制定, 其目标是土著数据的管理人和其他用户要“公平且关怀”。相比起2016年首次发表在《科学数据》上提出的“科学数据管理和管理的FAIR原则”, CARE原则是作为对于FAIR原则、以及其他开放数据运动的伦理补充。

全球开放科学数据状况

Dimensions的数据集专题库收录汇总了多个知名的科学数据平台的资源，对于跟踪了解当前全球的科学数据状况有一定的参考价值。截止2024年5月，该库共收录数据集31,507,869个。按机构所在国家来统计，TOP 5国家分别是美国（816,919条，占比2.6%）、中国（680,120条，占比2.2%）、德国（403,725条，占比1.3%）、英国（366,737条，占比1.2%）和法国（233,973条，占比0.7%）。

《开放数据现状报告》是Springer Nature每年联合多方机构发布的调研报告。2023年度报告显示，全球有超过6000名研究人员做出了回应，其中印度（12%）、中国（11%）和美国（9%）的回答比例最高。近四分之三的受访研究人员表示，他们仍然没有得到公开分享数据所需的支持。这些数据凸显了对加强社区合作和工具的需求增加，以支持研究人员转向可持续的开放科学实践。其余23%的受访者曾寻求并获得数据共享支持，支持主要来自内部来源（同事/主管 - 61%），其次是机构图书馆（31%）、研究办公室/内部机构专业知识（26%）、出版商（21%）和资助者（17%）。

全球开放科学数据的编目平台

科学数据仓储注册系统（re3data）：2015年5月，德国研究基金会（DFG）资助构建了re3data，是专门汇集全球不同学科科学数据仓储的注册平台，由德国洪堡大学柏林图书情报学院、德国波茨坦地学研究中心和德国卡尔斯鲁厄理工学院联合维护，2014年开始与DataBib合并，持续提供DataCite引用服务。re3data旨在对所有领域的科学数据仓储基于注册机制进行索引化和结构化描述，采用信息图标来描述每个科学数据仓储的基本特征，这样就能方便使用者更加方便、快捷地使用该系统。截止2024年5月，该系统共收录3,216条记录，其中中国的注册记录为，仅占比2.6%。

数据、标准和政策综合注册平台（FAIRSharing）：FAIRSharing的前身是2011年推出的Biosharing，通过注册操作，将标准（术语、格式和模型以及报告指南）、数据库和数据政策进行关联集成，并结合了所有三种类型资源的集成视图。目前共有1755条标准，2122条数据库（其中中国记录有126条，占比5.9%）和265条政策数据。

中国的国家科学数据中心

为了落实《科学数据管理办法》和《国家科技资源共享服务平台管理办法》的要求，规范管理国家科技资源共享服务平台，完善科技资源共享服务体系，推动科技资源向社会开放共享，2019年6月5日，科技部和财政部联合发布《国家科技资源共享服务平台优化调整名单》，发布了“国家高能物理科学

数据中心”等20个国家科学数据中心名单，具体如下。

序号	国家平台名称	依托单位	主管部门
1	国家高能物理科学数据中心	中国科学院高能物理研究所	中国科学院
2	国家基因组科学数据中心	中国科学院北京基因组研究所	中国科学院
3	国家微生物科学数据中心	中国科学院微生物研究所	中国科学院
4	国家空间科学数据中心	中国科学院国家空间科学中心	中国科学院
5	国家天文科学数据中心	中国科学院国家天文台	中国科学院
6	国家对地观测科学数据中心	中国科学院遥感与数字地球研究所	中国科学院
7	国家极地科学数据中心	中国极地研究中心	自然资源部
8	国家青藏高原科学数据中心	中国科学院青藏高原研究所	中国科学院
9	国家生态科学数据中心	中国科学院地理科学与资源研究所	中国科学院
10	国家材料腐蚀与防护科学数据中心	北京科技大学	教育部
11	国家冰川冻土沙漠科学数据中心	中国科学院寒区旱区环境与工程研究所	中国科学院
12	国家计量科学数据中心	中国计量科学研究院	市场监管总局
13	国家地球系统科学数据中心	中国科学院地理科学与资源研究所	中国科学院
14	国家人口健康科学数据中心	中国医学科学院	卫生健康委
15	国家基础学科公共科学数据中心	中国科学院计算机网络信息中心	中国科学院
16	国家农业科学数据中心	中国农业科学院农业信息研究所	农业农村部
17	国家林业和草原科学数据中心	中国林业科学研究院资源信息研究所	林草局
18	国家气象科学数据中心	国家气象信息中心	气象局
19	国家地震科学数据中心	中国地震台网中心	地震局
20	国家海洋科学数据中心	国家海洋信息中心	自然资源部

附录 5 开源合规案例精选

（1）中兴通讯开源合规案例

近十多年来，开源生态正变得越来越活跃和多样化，但是随着外部法规法条的日益完善以及市场准入门槛的提高，开源软件本身固有的合规、安全风险却长期困扰着绝大多数企业，特别是出海企业，成为企业开拓市场、合规经营头上的达摩克利斯之剑。

开源软件平均占到了整个软件产品成分的90%左右，开源软件所带来的风险也从一开始大家关注的开源许可证合规风险，逐步演进至EAR、产品安全、GDPR（个人信息保护）的合规和安全风险；近几年更是上升到了开源软件供应链安全、甚至是软件供应链安全的高度，开源软件的合规和安全也上升到各国的发展战略高度，成为企业之间、大国之间竞争的利器。在此背景下，构建适合企业发展战略的开源合规和安全治理体系就体现了企业开源治理新质生产力水平，在企业激烈的市场竞争中将起到举足轻重作用。

中兴通讯从2016年开始就尝试探索开源软件的合规治理，2018年出口管制事件后，公司全面重视合规，把合规上升到公司经营成功、成为百年老店目标的基石之一。在开源治理方面，由于开源软件的合规治理涉及法务、市场、研发、供应链、安全、研发流程和工具体系各个方面，究竟构建何种治理体系是经过了较长一段时间的碰撞、试点和反思。

中兴通讯以已有的高效产品开发（HPPD）研发体系为基础，将开源软件许可证、出口管制、产品安全、GDPR、软件供应链等合规、安全方面的外部法规、国际/国内标准、客户要求作为软件产品开发的约束条件，构建一体化的大合规、大安全的公司级治理体系，以确保能够满足外部不同法规、市场、客户的合规安全诉求。

治理模式	优势	劣势	生产力水平
项目级治理	快速响应，见效快	难以复制，共享难、综合成本最高	低
研究院级治理	较快响应，部分共享	跨院解决方案治理难	较低
公司级治理	共同守护，全共享，综合成本最优	范围广，流程、工具、能力要统一，时间长	最优

目前中兴通讯已经逐步构建起了公司级统一的开源合规和安全治理管控体系，通过信通院开源可信供应链和OpenChain双认证、对标欧盟弹性法案（CRA）和 BSI SBOM标准的NESAS 系列版本认证、GC 和 BSI BSZ认证、GSMA NESAS2.3（GSMA Network Equipment Security Assurance Scheme）移动通信设备产品开发和生命周期流程的独立安全审计、以及国内相关信创和国测方面的认证。

要想成功构建企业有效的开源治理新质生产力治理体系：

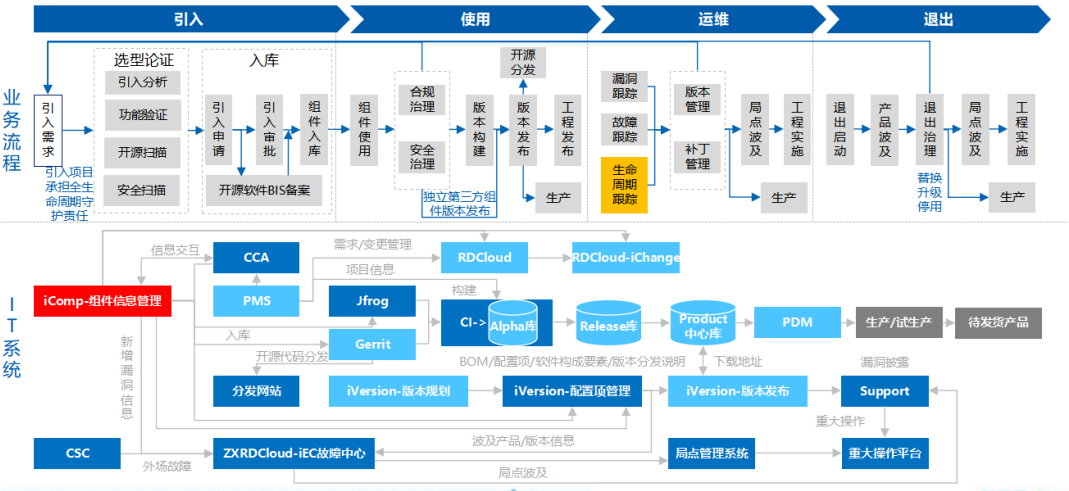
- 首先，公司领导对开源合规的重视是最关键的要素；
- 其次，全公司一盘棋，进行公司级的开源治理是成功的次要因素，也是提升开源治理新质生产力效能的最有效手段；
- 第三，企业须具备已固化的研发管控流程体系，开源合规和安全管控流程要求才能较容易的嵌入到已有的产品/项目研发管理过程中，从而确保管控的有效性；
- 第四，完整、全面、正确的识别产品/项目版本的软件成分（使用什么开源软件和如何使用开源软件？）是一切治理的基础！专业的SCA扫描工具是帮助项目了解开源现状和使用的必备工具；
- 第五，虽然是虚拟化的OSPO（开源软件管理办公室）运作，但是一个相对固定专业化的核心团队是确保治理方向正确、并引领企业逐年迭代演进的关键要素。

构建有效管控机制的过程并非一帆风顺，以下经验教训可供大家借鉴：

- 首先，组织层面对开源认知的不同，不了解开源的风险，随意用开源，害怕开源而不用开源是最可怕的。建议需要对组织各级进行洗脑、强有力的洗脑培训；
- 第二，组织高层领导的理解和支持是一切过程改进成功的必要条件，开源治理同样如此。外部合规安全事件驱动有时可起到事半功倍的效果。
- 第三、如何组织各领域专家、相关职能部门协同有序推进开源可信供应链建设？OSPO、开源COP建设、胡萝卜加大棒有时效果不错；
- 第四，如何说服组织内不同单位投入资源进行组织能力和工程建设，特别是组织一体化的开源治理管控体系？先试点成功，再全面推广，事实胜于雄辩；

● 第五，方向和目标非常重要，治理是一项长期的过程，需要说服领导和员工为什么要这样做？这样做有什么好处？组织的推进是帮助项目来进行合规和安全治理，否则这些工作只能项目自己来处理。解决了必须做的事情，剩下如何做就是你在帮助项目了；

开源软件和研发产品的全生命周期管理是开源治理管控的主线，自动化的研发、生产、运维系统是管控有效的必要保障，规范、流程、系统、度量展示与管控缺一不可。



开源可信供应链的要求没有尽头，企业需要持续优化、苦练内功、持续提升开源软件合规和安全管控新质生产力的研发效能，才能以不变应万变，在复杂的外部环境和激烈的市场竞争中立于不败之地。

(2) 麒麟软件开源合规案例

为保障使用软件包的来源合法，安全可信和开源合规，降低公司合规风险，麒麟软件公司特进行一系列开源合规相关工作，相关工作目前分为六大类，分别是引入合规，代码合规，安全可信，开源合规，健全制度与合规培训贯宣、理论支撑，后续也在持续优化中。

引入合规工作

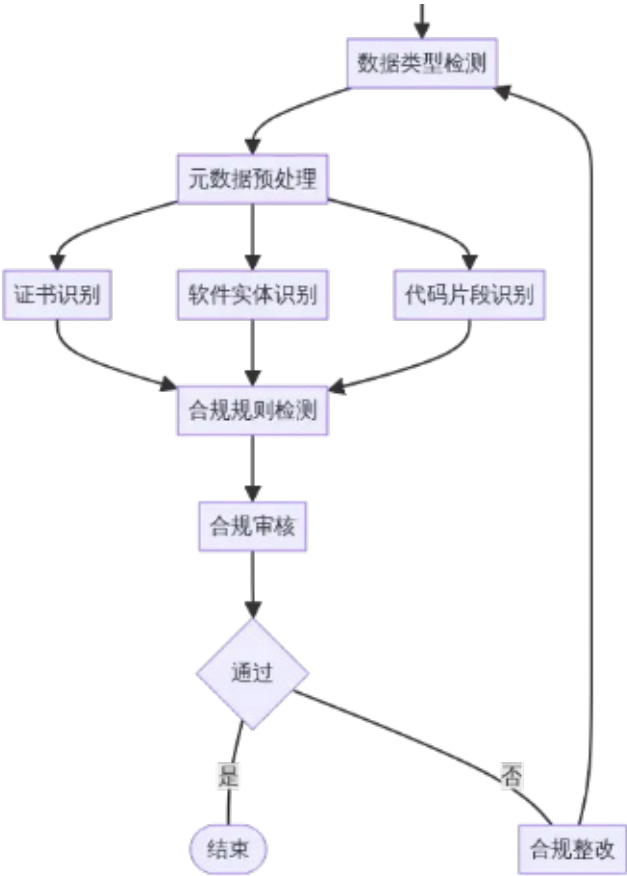
主要关注开源软件的引入选型。软件包中心仓是公司开展一切业务的资源起点，选型是外埠软件包引入中心仓的唯一途径。麒麟软件制定开源软件引入选型流程，规范了外埠软件包引入流程，做到

软件包生命周期管理端到端闭环；保障软件包的来源合法和安全可信；提供前期批量选型功能，提升研发效率。

引入原则基于技术、安全、合规、生命周期四个维度进行选型评估，确保最优引入，并根据选型相关信息，技术专家进行评审后才准许入库。

代码合规

麒麟软件搭建开源合规管理平台实现研发过程中对开源软件的引入、修改有基本的合规管理。开源合规管理平台的流程如下：



开源合规管理平台检测流程

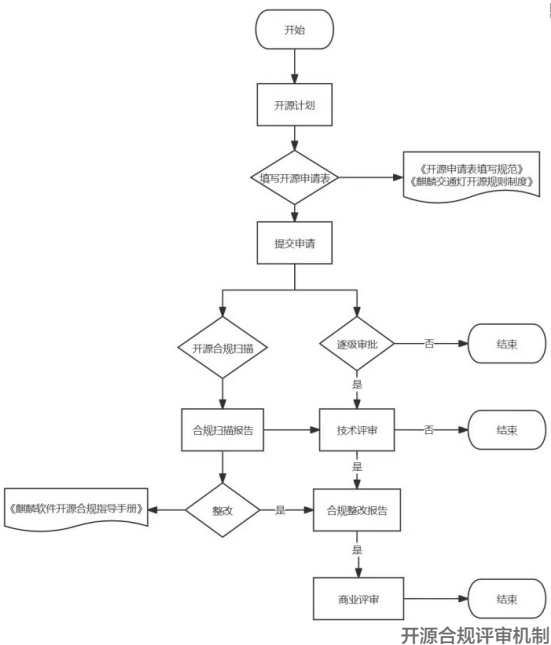
其中合规任务扫描支持软件包、URL地址、以及服务接入的方式创建合规扫描任务，合规平台先识别扫描任务的数据类型并进行相应的数据解析，再进行下一步的合规规则检测。合规规则检测的检测规则包括许可证兼容检测、版权规范检测、代码片段检测。合规审核为人工审核，当合规规则无法准确判断的合规问题的时候，会根据规则类型触发不同的审核流程，自动形成审核任务下发到软件包对应的责任人，审核后形成相应的记录存档在数据库。合规扫描完成后会形成一个完整的合规检测报告，覆盖合规检测问题以及标记具体位置，方便研发进行合规问题整改。

安全可靠

漏洞管理机制非常重要，其中漏洞发现与收集是指收集麒麟相关产品的安全漏洞信息，确保所有的安全漏洞问题能够得到闭环；漏洞分析是通过对产品的影响域、危害程度分析制定漏洞修复策略；漏洞修复是根据漏洞信息制定相关修复策略确保安全漏洞已经被彻底修复；漏洞维护与补丁披露是根据漏洞修复信息编写漏洞公告、披露并告知用户进行安全更新。

关于开源合规

为确保计划开源的代码合规开源成立了开源评审小组，通过合规扫描和整改流程，防范和解决潜在的代码合规风险，如版权侵权风险等，并从技术和商业角度评审，确保公司开源代码的合规性和价值。以下是麒麟软件的开源合规评审机制：



健全制度与合规培训贯宣

这项工作是麒麟软件开展合规工作的基础。开源软件的合规使用本质上属于知识产权的权利行使与处分相关问题，因此首先应制订企业整体知识产权管理制度，其中对于涉及开源事项的决策机构、管理模式及事项决策流程进行整体性规定。此外开源软件合规治理涉及企业人员众多，技术、产品、法务等多部门人员均会涉及开源软件合规治理节点，因此，针对新老员工、不同岗位职责设计不同的培训内容，通过多次培训将基本开源软件合规治理理念和知识传递给参与日常开源事项的每一名员工。

合规培训是指其中麒麟软件成立合规工作小组，主要负责规范制定、合规法律解读、代码开源及维护策略决策等。为明确开源软件管理的生命周期，定义各环节的执行规范而建立开源软件管理规范。

合规理论研究

麒麟软件法务部门对于构建企业开源软件合规治理体系提供坚实的理论支撑。具体表现为法务部门组织力量深入研究解读世界主要开源许可证的各项条款，仔细分析识别各开源许可证所规定的开源软件发布、分发、修改、使用，衍生等各种行为的法律义务，提炼合规要点，撰写各主要开源许可证合规使用指南文档为合规治理工作提供理论指导。

-  2-clause BSD许可协议合规指南240104
-  3-clause BSD合规使用指南1228
-  Apache 1.0合规使用指南 (BD-ZJX)
-  Apache 1.1合规使用指南(BD-ZJX0425)
-  Apache 2.0合规使用指南(1228)
-  CC-by知识共享协议
-  CDDL-1.0
-  EPL-2.0合规指南
-  GPLv2合规指南1229
-  MIT许可证合规指南
-  MPL2.0许可协议合规指南1
-  OFL字体许可证合规指南0108(1)
-  ZLIB许可证合规指南

指南文档

实时跟踪新发布开源许可证，对存量和增量开源许可证进行分类管理，建立不同研发场景的开源许可证准入原则，对不同开源许可证建立黑白灰名单。

法务部门时刻保持对国内外开源软件合规纠纷案例的分析与追踪，紧跟开源软件合规理论发展趋势，捕捉相关政策和司法裁判思路变化，及时调整企业合规治理政策以适应外部变化，对企业技术

部门日常落实开源软件合规治理中遇到的疑难问题进行排忧解难，全面护航企业开源软件合规治理工作的不断发展完善。建立完善的合规管理体系，是企业成功运营的重要前提和保障。

（3）统信软件开源治理案例

随着全球开源操作系统蓬勃发展，统信软件认为，当前头部操作系统公司规模呈现出“千百十”的特征，即：

- 千人以上的操作系统开发团队
- 百人以上的内核研发团队
- 十人以上的开源合规律师团队

这意味着，操作系统厂商对于开源治理愈发重视。而在开源治理方面，统信软件先人一步，引入开源合规工具，梳理开源合规流程，在产品立项、研发、发布等各个环节落地开展合规工作，具体包括：

- 在产品立项及研发过程中：遵守规范的第三方代码及依赖引入流程，经过SCA工具及人工审核以保证引入的第三方代码的合规性及与主程序的兼容性；
- 在项目开源工作中：遵守必要的代码通知规范，经过CI门禁及人工审核以保证开源通知的合规性及易读性；根据公司开源策略、项目开源目的及第三方代码/依赖的开源许可证情况综合考虑项目所选用的开源许可证；
- 在产品分发和发布工作中：遵守开源许可证的通知义务，经过人工审核以保证产品的分发材料满足开源许可证要求；
- 建立相应管理制度规范：为规范开源软件迁移流程，统信软件制定了《开源软件迁移控制程序》，确保多个环节协调及顺序工作的完整性，使迁移工作能够提高工作效率；
- 探索开源治理的评估方法与模型：制定了《统信UOS开源合规治理白皮书》，系统总结了统信UOS各版本的开源特性，阐述了主要开源许可证和重点开源项目的开源合规性审查的要点和方法。

企业的开源治理，包含了开源安全、开源合规、开源供应链等等，所以企业的开源合规工作也不是孤立的，往往需要与其他工作结合起来。

对于开源合规，主要集中在开源知识产权风险的应对上，如何遵守许可证规定的通知义务和开源义

务以避免触发著作权、专利、商标的侵权风险。此外，开源还涉及到域外管辖和出口管制的合规，例如美国出口管制条例EAR和我国的出口管制法规。当前开源许可证众多，每个许可证所规定的义务又各不相同，呈现出复杂多样、互不兼容的特点。这些都需要法务人员进行专业地分析和解读，而最后的合规处理又需要落实到代码和文件，需要由程序员来进行落地操作。

企业开源合规的成效，除了依赖于开源合规管理制度、开源合规工具，员工的开源素质是重点，包括员工对公司开源政策的理解，对重要许可证义务的理解，以及对标准合规处理的掌握。只有不断通过培训提高员工的开源素质，开源合规工作才能事半功倍。

开源合规工作的难点，应该是企业不断增长的合规需求与合规工具有限功能之间的矛盾。开源合规需求很多，包括：软件物料清单、组件识别、篡改识别、代码溯源、依赖检测、许可证兼容性分析等等，各种合规工具在不同需求点上表现不一；当前，由于开源供应链的合规需求上升，对软件物料清单的准确性和完整性又提出了更高的要求。

企业建立适合自身的开源政策最为重要。有的企业只是在企业内部使用开源软件，并不做分发，它要遵守的开源合规义务就很少，主要需要考虑的可能就是开源供应链上游带来的合规风险；有的企业在产品中使用开源软件也做分发，但对自研代码要求闭源，那就要高度关注开源代码和开源组件的选择和使用方式，避免违反copyleft许可证带来的开源义务；有的企业是既使用开源也向外开源，既要避免许可证冲突又要符合开源通知义务。企业必须根据自身战略和定位制订清晰的开源政策，并以其指导开源合规管理办法的制订，以及开源合规管理组织的职责和工作。

关于开源合规方面的两点建议：

跟所有其他的企业合规工作一样，开源合规也需要合规成本，可能是人力成本，也有可能是产品延迟发布的时间成本，所以开源合规的开展应以对项目最小影响为目标，科学地设置审核节点和审核内容非常重要。例如：在立项阶段就要求研发经理向合规审核小组提出开源使用请求，包括：项目是否开源、项目拟采用的开源许可证、项目拟采用的组件或第三方库及调用方式等，避免事后因重大合规缺陷导致的返工。还有可以在第一轮全量测试时进行项目代码审计，既能对项目代码进行完整扫描，发现问题又能即时返回研发进行代码修改。

行动起来最重要，不用等到所有资源全部到位，也不用等到各项功能都完全达标的合规工具，开源合规审核最重要的因素还是人。哪怕是简易审核流程，也能将最重要的不合规项发现出来。

致谢

《2024中国开源发展现状》由COPU牵头，联合中国开发者社区CSDN、中国科学院软件研究所、北京开源创新委员会等80多家企业及行业机构、102位开源专家和志愿者共同编撰完成，写作过程中得到许多海内外开源界人士、企业单位、开源组织、社区、科研机构、高等院校的大力支持，在此表示感谢！

编委会

顾问

陆首群 COPU名誉主席

策划组

刘 澎 COPU副主席兼秘书长，北京开源创新委员会主任

赵 琛 中国科学院软件研究所所长

蒋 涛 CSDN创始人&董事长，开源代码托管平台GitCode发起人，COPU副主席

梁志辉 COPU常务副秘书长

主编组

孟迎霞 COPU常务副秘书长，CSDN副总裁

宋可为 北京开源创新委员会常务副主任，COPU常务副秘书长

隆云滔 中国科学院科技战略咨询研究院副研究员

王 伟 华东师范大学数据科学与工程学院教授

孙 启 北京开源创新委员会副秘书长、北京长风信息技术产业联盟副秘书长

荆 琦 北京大学副教授，COPU副秘书长

谭中意 开放原子基金会TOC副主席，COPU常务副秘书长

唐小引 CSDN开源总编，《新程序员》执行总编

鞠东颖 COPU执行副秘书长

工作组 (按首字母排序)

- 安 泐 Circle Linux社区架构师
- 陈 伟 COPU专家委员会副主任
- 程智锋 北京地平线信息技术有限公司总监
- 邓 伟 英特尔亚太研发有限公司AI软件布道师
- 丁禹民 清华大学公共管理学院研究员
- 韩凡宇 华东师范大学博士生
- 何 苗 CSDN开源编辑
- 贾朔维 中国网络空间研究院信息化研究所研究员
- 梁华博 山西农业大学软件工程专业学生
- 李 鑫 CSDN UI设计师
- 刘明明 北京初心使命软件有限公司经理
- 刘巍巍 北京开源创新委员会成员、北京初心使命软件有限公司经理

贡献者 (按首字母排序)

- 包云岗 中国科学院计算技术研究所副所长、研究员
- 白国华 中国开源软件推进联盟PostgreSQL分会秘书长、北京象前行信息科技有限公司总经理
- 边思康 蚂蚁集团开源办公室负责人
- 常 青 阿波罗智能技术(北京)有限公司生态部产品运营经理
- 陈 绪 阿里云计算有限公司技术战略总监
- 陈 渝 清华大学计算机系软件所副所长
- 陈梓立 Apache软件基金会成员
- 陈阳 开源社联合创始人、理事长
- 丁紫薇 麒麟开源合规开发工程
- 都莉楠 北京赛迪出版传媒有限公司副总理、软件和集成电路杂志社社长

- 宁 固 COPU常务副秘书长
- 宋逸群 阿里云战略发展部北京政策研究中心主任
- 滕召智 OpenSDV汽车开源联盟总监
- 屠 敏 CSDN业界主编
- 王蕴博 北京字节跳动科技有限公司首席开源布道师
- 许哲平 中国科学院文献情报中心副研究馆员
- 袁 新 中国网络空间研究院信息化研究所研究员
- 张红月 CSDN运营
- 郑丽媛 CSDN编辑
- 章文嵩 AutoMQ 联合创始人&首席战略官, 开源项目LVS创始人

- 冯 铃 清华大学计算机学院教授
- 冯 媛 IBM架构师
- 顾雨婷 统信软件产品行销经理
- 郭磊涛 重庆中科汽车软件创新中心主任
- 郭鑫宇 中国兵器207所处长
- 郭振宇 腾讯云副总裁
- 何小庆 嵌入式系统协会副理事长, 嵌入式操作系统专家
- 胡 旷 阿波罗智能技术(北京)有限公司Apollo布道师
- 胡欣慰 openEuler社区技术委员会主席
- 华国栋 江苏省智能网联汽车产业创新联盟理事长
- 贾 伟 北京智谱华章科技有限公司技术运营总监

姜 宁	Apache软件基金会董事	王 婧	涛思数据联合创始人
康艳红	开源社区SIG负责人	吴国斌	北京嘀嘀无限科技发展有限公司总监
Keith Chen	Linux基金会亚太区经理	武会添	重庆中科汽车软件创新中心科研管理部副部长
李昊阳	开源基础设施基金会中国区总经理	谢宝友	国科础石总工程师
李 蒙	中国移动物联网处长	项曙明	中兴通讯开源合规&安全治理总监
李暮月	中国电科普华基础软件经理	熊谱翔	睿赛德总经理
李小平	中科方德经理	徐贵洲	翼辉信息副总经理
李莹烛	重庆中科汽车软件创新中心副主任	徐立锋	中兴通讯操作系统产品部总工程师
李运喜	中国航空631所处长	杨 聪	苏州大学未来科学与工程学院副教授
李震宁	麒麟软件副总裁	杨丽蕴	中国电子技术标准化研究院研究室主任 中国开源云联盟常务副秘书长
梁 尧	OpenSDV汽车开源联盟副总监	姚 勇	IBM云平台研发总监
林旅强	零一万物开源负责人	于仕琪	南方科技大学教授
刘大力	IBM资深软件工程师	袁 恽	IBM 软件工程师
刘 睿	IBM资深软件工程师	袁 东	vivo经理
马 诺	零一万物开发者关系工程师	张家驹	红帽大中华区首席架构师
马 涛	龙蜥社区理事长, 阿里巴巴开源技术委员会副主席	张诗明	OPPO开源经理
孟繁晶	IBM中国系统开发中心CTO	张嗣元	IBM资深开发经理
隋 伟	北京萝卜的壳科技有限公司算法总监	赵海玲	开放原子开源基金会行研部
王 蓉	阿里巴巴开源办公室运营经理	赵 婕	小米技术委员会经理
王 峰	上海安势信息技术有限公司总经理	翟艳芬	中国软件评测中心副主任
王帅峰	北京初心开源技术有限公司经理	钟卫东	中瓴智行首席科学家
王耀华	统信软件开源社区中心总经理、深度社区负责人		

支持单位

开源社区 (按首字母排序)

CSDN开发者社区	GitCode 代码托管平台	Circle Linux社区
龙蜥社区	Gitee 代码托管平台	开源社
openEuler社区	Gitlink代码托管平台	X-lab开放实验室
AtomGit 代码托管平台	木兰开源社区	

企事业单位（按首字母排序）

阿波罗智能技术（北京）有限公司	北京字节跳动科技有限公司	小米科技有限责任公司
阿里云（北京）科技有限公司	重庆中科汽车软件创新中心	翼辉信息技术有限公司
北京奥星贝斯科技有限公司	重庆开源共创科技有限公司	英特尔亚太研发有限公司
北京初心开源技术有限公司	国际商业机器公司（IBM）	浙江九州未来信息科技有限公司
北京创新乐知网络技术有限公司（CSDN）	国科础石（重庆）软件有限公司	中国兵器工业集团有限公司
北京地平线信息技术有限公司	红帽（Red Hat）	中国电子科技集团有限公司
北京嘀嗒无限科技发展有限公司	京东集团股份有限公司	中国航空工业集团有限公司
北京萝卜的壳科技有限公司	零一万物	中国科学院计算技术研究所
北京平凯星辰科技发展有限公司	蚂蚁科技集团股份有限公司	中国联合网络通信集团有限公司
北京思斐软件技术有限公司	麒麟软件有限公司	中国软件评测中心
北京涛思数据科技有限公司	睿赛德科技（上海）有限公司	中国移动通信集团有限公司
北京易捷思达科技发展有限公司	上海安势信息技术有限公司	中科方德软件有限公司
北京智谱华章科技有限公司	深圳前海微众银行股份有限公司	中瓴智行（成都）科技有限公司
	腾讯科技（北京）有限公司统信软件技术有限公司	中兴通讯股份有限公司
	维沃移动通信有限公司	OPPO 广东移动通信有限公司
		北京涛思数据科技有限公司

基金会

Apache软件基金会	开放原子开源基金会	天工开物开源基金会
Linux基金会亚太区	开源基础设施基金会中国区	

行业组织

COPU	北京智源人工智能研究院	中国嵌入式系统协会
OpenSDV汽车开源联盟	江苏省智能网联汽车产业创新联盟	北京开源创新委员会
北京开源芯片研究院		

学术机构

北京大学	南方科技大学	中国科学院科技战略咨询研究院
北京航空航天大学	清华大学	中国科学院文献情报中心
华东师范大学	苏州大学未来科学与工程学院	

2024 中国开源发展现状

2024 China Open Source Development Status



COPU·开源联盟



扫码下载电子版



扫码参与内容共建

COPU·开源联盟

秘书处电话: +86 10-88558999

联盟邮箱: office@copu.org.cn

联盟官网: <http://www.copu.org.cn>

地址: 北京市海淀区紫竹院路66号