

科学研究动态监测快报

信息科技专辑

INFORMATION TECHNOLOGY MONTHLY EXPRESS

2025
第5期

总第257期

本期视点

美斯坦福大学发布《2025人工智能指数报告》

美无限期限制英伟达H20芯片对华出口

欧盟发布《人工智能大陆行动计划》

英国政府投资1.21亿英镑用于量子技术开发

美国研究人员开发低功耗RISC-V AI芯片

NVIDIA计划开发AI超级计算机

中国科学院成都文献情报中心

CHENGDU LIBRARY AND INFORMATION CENTER, CHINESE ACADEMY OF SCIENCES

目 录

重点关注

[信息技术]美斯坦福大学发布《2025 人工智能指数报告》 ..1

科技政策与科研计划

[半导体]美无限期限制英伟达 H20 芯片对华出口3

[人工智能]欧盟委员会征求通用人工智能模型规则意见3

[人工智能]欧盟发布《人工智能大陆行动计划》4

[人工智能]欧盟将投资 13 亿欧元发展网络安全和 AI 技术...5

[量子技术]欧盟发布《量子技术能力框架 3.0 版本》6

[量子技术]英国政府投资 1.21 亿英镑用于量子技术开发6

[网络安全]瑞典发布《2025-2029 年国家网络安全战略》7

前沿研究动态

[RISC-V]美国研究人员开发低功耗 RISC-V AI 芯片8

[半导体]台积电发布晶圆级封装技术 SoW-X9

产业动态

[人工智能]NVIDIA 计划开发 AI 超级计算机.....9

[人工智能]OpenAI 宣布开发开源 AI 模型10

[RISC-V]香山高性能处理器设计融入主流编译器基础设施 11

执行主编：唐川

执行编辑：黄茹

E-mail: tangc@clas.ac.cn

E-mail: huangr@clas.ac.cn

出版日期：2025 年 5 月 1 日

重点关注

[信息技术]美斯坦福大学发布《2025 人工智能指数报告》

2025年4月7日，美国斯坦福大学发布第八版《2025人工智能指数报告》，揭示了2024年人工智能行业的12个关键趋势。报告指出，2024年人工智能在技术能力、资金投入及政策监管方面均达到历史新高。

（1）人工智能系统性能持续提升

2024年人工智能系统在MMMU、GPQA以及SWE-bench三大基准上的得分较2023年分别提升18.8%、48.9%和67.3%。此外，在生成高质量视频方面，语言模型智能体在有限时间的编程任务中超越了人类。

（2）人工智能正迅速从实验室走向日常生活

2023年美国食品药品监督管理局(FDA)批准了223种人工智能医疗设备，而2015年仅有6种。美国自动驾驶汽车公司Waymo每周提供的自动驾驶出行超过15万次，百度“萝卜快跑”自动驾驶出租车车队现已在中国多个城市提供服务。

（3）人工智能的投资和使用创历史新高

2024年，美国私营部门在人工智能领域的投资总额达到1091亿美元，几乎是同期中国（93亿美元）的12倍、英国（45亿美元）的24倍。其中，生成式人工智能表现尤为突出，2024年全球范围内吸引的私营投资总额高达339亿美元，较上一年增长18.7%。同时，人工智能的商业化应用加速推进。2024年，有78%的组织采用了人工智能技术，较前一年增长55%。此外，越来越多的研究表明，人工智能不仅提高了生产力，而且在多数情况下，还有助于缩小劳动力市场中的技能差距。

（4）中国与美国在顶级人工智能模型开发方面差距缩小

2024年，美国机构推出了40个引人注目的人工智能模型，显著超过中国的15个和欧洲的3个。尽管美国在模型数量上保持领先，但在模型性能方面中国与美国之间的差距迅速缩小。在MMLU和HumanEval等主要基准测试中，两者间的性能差异已从2023年的两位数百分比缩小到2024年的近乎相等。同时，中国在人工智能领域的出版物和专利数量方面继续保持领先。此外，模型开发呈现出越来越全球化的趋势，中东、拉丁美洲和东南亚等地区也纷纷推出了新模型。

（5）负责任的人工智能生态系统呈现出不均衡的发展态势

目前，针对负责任的人工智能系统的标准化评估较为稀缺。在企业层面，识别人工智能风险与采取的实质行动之间存在差距。但在政府层面，人工智能生态系统的治理不断加强。2024年，全球在人工智能治理方面的合作得到加强，经济合作与发展组织、欧盟、联合国和非洲联盟在内的多个国际组织发布了聚焦于透明度、可信度以及其他负责任人工智能原则的框架。

（6）全球对人工智能的乐观情绪呈上升趋势，但地区间仍存在分歧

在中国、印度尼西亚和泰国，大多数民众认为人工智能产品和服务带来的益处大于潜在的弊端，2024年持乐观态度的占比分别为83%、80%、77%。加拿大、美国和荷兰则正好相反，2024年持乐观态度的占比分别为40%、39%、36%。总体而言，自2022年以来，多数国家对人工智能持乐观态度的比例呈现上升趋势。截至2024年，德国的乐观情绪增长了10%，法国增长了10%，加拿大增长了8%，英国增长了8%，美国增长了4%。

（7）人工智能正变得愈发高效、经济实惠且易于获取

在2022年11月至2024年10月期间，达到GPT-3.5水平的人工智能系统推理成本下降了280倍。从硬件层面来看，成本每年降低30%，能效每年提升40%。同时，开放权重模型与封闭模型之间的性能差距也在不断缩小，在某些基准测试中，一年内性能差异从8%降至仅1.7%。这些趋势共同作用，迅速降低了高级人工智能的准入门槛。

（8）各国政府不断加强对人工智能的监管和投资力度

2024年，美国联邦机构出台了59项与人工智能相关的法规，是2023年的两倍多，且发布法规的机构数量也达到了2023年的两倍。在全球范围内，自2023年以来，提及人工智能立法的国家数量增加了21.3%，自2016年以来，这一数字增长了9倍。同时，各国政府也不断加大对人工智能的投资：加拿大承诺投入24亿美元，法国承诺投入1090亿欧元，印度承诺投入12.5亿美元，沙特阿拉伯则发起了一项1000亿美元的倡议，中国也启动了475亿美元的半导体基金。

（9）人工智能和计算机科学教育正在扩大，但普及程度仍不够

目前，已有三分之二的国家提供或计划提供K-12计算机科学教育，这一比例是2019年的两倍。在美国，计算机专业学士学位的毕业生人数在过去十年中增长了22%，81%的K-12计算机科学教师认为人工智能应成为计算机科学基础教育的一部分，但只有不到一半的教师认为其具备教授人工智能的能力。

（10）工业界人工智能模型研发依然领先

2024年，近90%的先进人工智能模型来自产业界，高于2023年的60%，而学术界仍然是高引用率研究的首要来源。然而，模型间性能差距正在缩小：排名第一位和第十位模型间的得分差距在一年内从11.9%降至5.4%，排名前两位模型间的得分差距现在仅为0.7%。

（11）人工智能在科学研究领域的重要性日益凸显

2024年，诺贝尔物理学奖和化学奖分别表彰了人工智能在深度学习和蛋白质折叠中的应用，而图灵奖则表彰了人工智能在强化学习方面的开创性贡献。

（12）复杂推理仍是一项挑战

尽管人工智能模型在诸如国际数学奥林匹克竞赛问题等任务中表现出色，但在

PlanBench等复杂推理基准测试中，仍然面临诸多挑战。即使存在可证明正确的解决方案，这些模型也常常无法可靠地解决逻辑任务，在一定程度上限制了模型在高风险环境中的有效性。

唐衢 黄茹 供稿自

<https://hai.stanford.edu/ai-index/2025-ai-index-report>

原文标题：Stanford HAI's 2025 AI Index Reveals Record Growth in AI Capabilities, Investment, and Regulation.

科技政策与科研计划

[半导体]美无限期限制英伟达 H20 芯片对华出口

2025年4月14日，美国政府以H20芯片可能用于中国超级计算机且存在潜在军事用途为由，宣布无限期限制英伟达H20芯片对华出口，未来任何相关出货均需提前获取特殊许可证。

H20芯片是英伟达为符合美国此前出口管制规定专门设计的芯片，相比其旗舰产品H100性能有所降低。H20芯片配备了96GB的HBM3内存，运行速度高达4TB/s，提供约296万亿次混合精度计算能力，每个芯片的性能密度约为2.9万亿次。

据彭博情报分析师预计，若限制措施持续，英伟达与中国相关的收入占总销售额的比例，可能从2024财年的约13%降至个位数的低位或中位。随着客户寻求替代推理加速器，华为等中国人工智能企业有望受益。

李星元 黄茹 供稿自

<https://www.techpowerup.com/335534/us-bans-export-of-nvidia-h20-accelerators-to-china-a-potential-usd-5-5-billion-loss-for-nvidia#:~:text=The%20Commerce%20Department%20delivered%20the%20news%20to%20NVIDIA,redirected%20into%20Chinese%20supercomputers%20with%20potential%20military%20applications>

原文标题：US Bans Export of NVIDIA H20 Accelerators to China, a Potential \$5.5 Billion Loss for NVIDIA

原文标题：US Bans Export of NVIDIA H20 Accelerators to China, a Potential \$5.5 Billion Loss for NVIDIA

[人工智能]欧盟委员会征求通用人工智能模型规则意见

2025年4月22日，欧盟委员会邀请利益相关方为通用人工智能模型制定规则，旨在为即将出台的欧盟委员会指南提供支持，该指南将解释《人工智能法案》中有关通用人工智能模型条款的关键概念，明确界定“通用人工智能模型”“供应商”“市场投放行为”等核心范畴。

此次意见征集面向通用人工智能模型供应商、人工智能系统下游供应商、民间团体、学术界、专家及公共部门等，要求各方于2025年5月22日前反馈。该指南还将细化人工智能办公室出台的相关举措，并补充目前正在定稿的《通用人工智能行为准则》。若《通用人工智能行为准则》获欧盟委员会人工智能办公室和人工智能委员会批准，供应商签署后可通过该准则减轻行政负担，并将其作为合规运营的基准。

该指南及最终版《通用人工智能行为准则》预计于2025年8月前公布。此外，欧盟委员会还计划针对人工智能系统高风险类别认定开展协商，助力利益相关方落实《人工智能法案》。

李星元 黄茹 供稿自

<https://digital-strategy.ec.europa.eu/en/news/commission-seeks-input-clarify-rules-general-purpose-ai-models>

原文标题：Commission seeks input to clarify rules for general-purpose AI models

[人工智能]欧盟发布《人工智能大陆行动计划》

2025年4月9日，欧盟委员会发布了《人工智能大陆行动计划》，立足于欧盟独特的产业生态和制度优势，基于“五大战略支柱”，全面提升欧盟在人工智能领域的全球竞争力。

（1）强化人工智能计算基础设施建设

欧盟委员会计划建立覆盖全欧洲的人工智能工厂网络，具体包括：依托欧盟领先的超级计算网络，在欧盟至少建立13个AI工厂，为初创企业、研究机构和传统产业提供尖端人工智能开发平台；筹建5个“AI超级工厂”，专注前沿模型训练。此外，欧盟委员会还计划推出《云计算和人工智能发展法案》，通过政策引导促进私人投资，目标是在未来5至7年内使欧盟数据中心容量至少增长两倍，并优先建设高能效的可持续数据中心。

（2）增强高质量数据开发能力

高质量数据资源是人工智能发展的核心基础，欧盟委员会计划实施双轨并行的数据战略：一方面出台《数据联盟战略》，构建欧盟统一的内部数据市场，打通数据流通壁垒；另一方面在人工智能工厂设立专业数据实验室，收集整理不同来源的高质量数据，提升欧盟数据资源的可利用性和质量。

（3）推动人工智能在战略领域的应用

目前，仅有13.5%的欧盟企业应用人工智能，欧盟委员会计划启动“人工智能应用战略”，重点推动医疗、公共管理等关键领域的人工智能应用。据悉，该战略将充分利用人工智能工厂、欧洲数字创新中心等人工智能创新基础设施，加速人工智

能技术在各行业的落地实施。

（4）加快人工智能专业人才引进和培养

为满足欧盟日益增长的人工智能人才需求，欧盟委员会一方面会继续优化高技能人工智能人才引进政策，促进国际招聘，为优秀人工智能人才提供合法移民途径；另一方面欧盟委员会将继续加强人才培养，在关键领域推进人工智能教育与培训项目，培养下一代人工智能专家，并支持劳动者提升技能和进行技能再培训。

（5）简化监管

欧盟委员会正全力推进《人工智能法案》的落地实施，通过发布配套实施细则和操作指南、制定行业最佳实践标准等举措，为企业提供清晰的合规路径。同时，欧盟委员会即将设立专门的一站式咨询服务平台，为企业提供资讯和合规指导。

唐衡 供稿自

https://commission.europa.eu/topics/eu-competitiveness/ai-continent_en

原文标题：Shaping Europe' s leadership in artificial intelligence with the AI continent action plan

[人工智能]欧盟将投资 13 亿欧元发展网络安全和 AI 技术

2025 年 3 月 28 日，欧盟委员会发布 2025-2027 年数字欧洲计划（DIGITAL），拟拨款 13 亿欧元用于部署对欧洲未来和技术主权具有重要战略意义的关键技术。该工作计划侧重于人工智能的部署及其在企业 and 公共管理、云计算、网络韧性和数字技能方面的应用。欧洲战略技术平台（STEP）也将通过加速创新，为有发展前景的项目提供 STEP 质量标签，以增加其获得公共和私人投资的机会。

DIGITAL 的重点优先事项包括：

（1）提高生成式人工智能应用程序的可用性和可访问性。资助资金将用于测试“虚拟世界”的沉浸式环境、实施《人工智能法案》及部署节能型公共数据空间。该措施对建设人工智能工厂、为企业和公共部门开发生成式人工智能模型至关重要。

（2）支持欧洲数字创新中心（EDIHs）为企业和公共机构提供专业技术知识、技术测试和技能辅导，以推动人工智能在欧洲企业和公共机构中的普及。

（3）支持“目标地球”（DestinE）计划，致力于构建地球数字模型，以支持气候适应和灾害风险管理。资助资金将用于增强地球数字模型的功能，让更多的研究人员可访问。

（4）增强网络韧性。通过实施欧盟网络安全储备等网络安全解决方案，提高包括医院和海底电缆在内的关键基础设施的韧性和安全性。

（5）提升欧盟教育和培训机构的数字技能培养能力，增强人才的高级技能。

（6）促进欧盟“数字身份钱包计划”（EUDI Wallet）和欧洲信托基础设施在成

员国的使用。

(7) 构建高效、优质、互操作性强的数字公共服务体系，推动公共机构转型。

唐衡 供稿自

https://ec.europa.eu/commission/presscorner/detail/en/ip_25_907

原文标题: Commission to invest € 1.3 billion in artificial intelligence, cybersecurity and digital skills

[量子技术]欧盟发布《量子技术能力框架 3.0 版本》

2025年4月15日，欧盟推出了量子技术能力框架（CFQT）3.0版本及配套的认证方案，为量子技术领域能力评估与人才发展提供指引。

量子技术能力框架是欧洲量子技术能力与资质的参考标准。该框架涵盖知识技能图谱，系统梳理了量子技术核心知识体系与技能要求，划分三大专业领域（如量子计算、量子通信等），每个领域设6级熟练度，并定义九种岗位所需资质要求，可用于规划教育活动、评估个人资质与岗位要求。最新的3.0版本，主要是在2.5版本基础上进行了一些细微的完善。

新认证方案同步上线，通过样例任务与考试可以量化量子技术能力熟练度。量子技术能力框架应用广泛，教育者可用以优化课程设计；学生能据此评估自身水平、定制学习目标；招聘方得以精准定位人才、优化招聘；专业人士和求职者可规划职业路径、提升竞争力，助力欧洲量子技术人才生态建设。

李星元 黄茹 供稿自

https://qt.eu/news/2025/2025-04-15_release-of-competence-framework-for-qt-version-3.0-and-additional-certification-scheme.php

原文标题: Release of Competence Framework for Quantum Technologies version 3.0 and additional certification scheme

[量子技术]英国政府投资 1.21 亿英镑用于量子技术开发

2025年4月14日，英国政府宣布将投资1.21亿英镑用于发展量子技术，旨在巩固英国在该领域的全球领先地位，并推动经济增长。

这笔资金在未来一年内的投资情况如下：

(1) 由英国创新署负责筹集4610万英镑，用于加速量子技术在计算、网络、定位导航与授时（PNT）、传感等多个领域的实际应用；

(2) 向国家量子计算中心投资2100万英镑，推进其试验平台项目，同时获得量

子软件实验室助力，探索量子技术革新路径；

(3) 向国家物理实验室的量子测量计划投资1090万英镑，鼓励企业深度利用量子技术；

(4) 向2024年设立的5个研究中心投资2360万英镑，其中300万专门用于培训和技能培养；

(5) 1510万英镑用于11个量子技术奖学金项目，挖掘量子技术在药物研发、疾病诊断等更多现实场景的应用；

(6) 430万英镑用于支持早期职业研究人员和量子相关人才培养项目。

量子技术凭借对宇宙最小粒子特性的运用，可构建超强计算机与传感器。英国此举不仅助力打击欺诈、反洗钱等犯罪行为，还将在医疗保健、电网能源效率提升等领域带来利好。此外，投资也为年轻研究人员提供了创新平台，有望催生新企业和就业岗位，推动英国经济发展。

目前，英国量子领域发展势头强劲，拥有全球第二大量子企业群体，仅次于美国。此次投资是英国政府变革计划的重要一环，基于诸如《人工智能机遇行动计划》等政策，进一步释放量子技术潜力，助力英国在全球科技竞争中抢占先机。

黄茹 李星元 供稿自

<https://www.gov.uk/government/news/121-million-boost-for-quantum-technology-set-to-tackle-fraud-prevent-money-laundering-and-drive-growth>

原文标题：£121 million boost for quantum technology set to tackle fraud, prevent money laundering and drive growth

[网络安全]瑞典发布《2025-2029 年国家网络安全战略》

2025年3月20日，瑞典政府正式向议会提交了《2025-2029年国家网络安全战略》，对未来五年的国家网络安全建设进行了整体谋划，以应对日益复杂的网络安全威胁，提升国家整体网络安全水平。该战略是瑞典落实欧盟《网络与信息安全指令(NIS2)》的一项重要举措，提出了以下三大支柱性措施。

(1) 构建高效网络安全工作体系

瑞典政府将采取的具体措施包括：推动公私部门开展系统有效的网络安全实践，强化对关键基础设施的保护，防范供应链相关风险。要求各行业部门实施风险管理和相应的网络安全防护措施，尤其是NIS2指令中涉及的关键领域；提升公共部门的网络安全能力，统一相关网络安全标准；通过立法进一步加强供应链安全，明确采购流程，减少对单一供应商的依赖；简化优化相关法规实施，推动各部门之间的协调监管；为中小企业提供必要支持，帮助其提升网络安全能力和水平。

（2）提升网络安全知识与能力水平

该措施旨在增强全社会的网络安全意识，填补相关人才缺口，促进研发创新。瑞典政府将采取的具体措施包括：加强网络安全公共宣传，推广“网络卫生”概念，提升公众网络安全意识；将网络安全嵌入基础教育及高等教育课程，创设“瑞典网络校园（Cybercampus Sverige）”，培养网络安全专门人才；加强公私领域合作，重点支持AI、量子技术等前沿技术在网络安全方面的应用；积极参与国际标准制定，确保瑞典在加密技术方面的领先地位。

（3）强化网络安全事件预防处置机制

瑞典政府将采取的具体措施包括：完善网络安全事件响应机制，将瑞典国家网络安全中心（NCSC）确立为网络安全核心枢纽，优化网络安全事件报告及跨部门协作流程；通过安全演习、实时信息共享等措施提升关键行业领域的应急能力；增强国内执法部门打击网络犯罪的能力，推动网络执法跨国合作。

除上述三大战略支柱外，该战略还明确了当前面临的重大挑战及相应的应对举措，具体包括：针对敌对国家威胁，进一步提升关键基础设施的网络安全防护能力和主动防御水平；针对网络犯罪，重点打击勒索软件攻击及数据窃取犯罪活动，为执法部门配备更丰富的技术工具；针对供应链安全，要求相关企业对供应商风险进行评估，并贯彻落实欧盟《网络弹性法案》的相关要求；针对技术可控性，进一步降低对非欧盟技术的依赖程度，发展本土加密技术；针对人才缺口，强化网络安全教育体系，拓展职业培训项目，并积极开展国际合作。

唐衡 供稿自

<https://www.regeringen.se/pressmeddelanden/2025/03/carl-oskar-bohlin-presenterade-en-ny-nationell-strategi-for-cybersakerhet/>

<https://www.regeringen.se/contentassets/0903061f79204084b6acf4ce1a978830/nationell-strategi-for-cybersakerhet-2025-2029.pdf>

原文标题：Carl-Oskar Bohlin presenterade en ny nationell strategi för cybersäkerhet

前沿研究动态

[RISC-V]美国研究人员开发低功耗 RISC-V AI 芯片

2025年4月21日，美国加州大学研究人员开发了一款低功耗多核AI芯片——MAVERIC，这款芯片专为机器人应用打造，采用16nm工艺制造，集成了4个RISC-V CPU核心和13个人工智能加速器核心，在机器学习和机器人任务中表现出色，其人工智能视觉分析帧率高达72帧/秒。

MAVERIC芯片的一大核心优势是支持机器人重建周边环境，通过融合深度估计

(DE)与同步定位和映射(SLAM)技术,显著提升感知能力。在东京超大规模集成电路(VLSI)技术与电路研讨会上展示的数据显示,该芯片运行速度最高可达1GHz,峰值能效达8TOPS/W。该芯片还支持闭环操作,在端到端融合深度估计与同步定位和映射过程中,每帧能耗10mJ,帧率72帧/秒。

芯片硬件配置同样亮眼,配备8个8位整数加速器、5个32位浮点加速器,供电电压0.52V-1.0V,100MHz时功耗低至20mW,1GHz运行时峰值功耗1.8W。

李星元 黄茹 供稿自

<https://hardwarebee.com/electronic-breaking-news/low-power-ai-chip-for-robotics/#:~:text=The%20chip%20boasts%20a%2016nm%20process%20and%20features,per%20second%20for%20machine%20learning%20and%20robotics%20tasks.>

原文标题: Low power AI chip for robotics

[半导体]台积电发布晶圆级封装技术 SoW-X

2025年4月24日,中国台湾地区台积电公司在近期举办的2025年北美技术论坛上,发布晶圆级封装技术SoW-X,计划于2027年实现量产,届时将可提供高达9.5倍的光罩尺寸。

台积电推出了其CoWoS封装技术的升级版SoW-X。与上一版SoW-P不同,SoW-X采用“中介层先行”的后芯片安装流程,即先在晶圆上构建中介层,然后再添加芯片。相比于传统方案,SoW-X方案的光罩面积将达到7722平方毫米,显著提升了有限空间内的芯片集成度和处理能力。对于高性能计算、人工智能等领域,更大的封装空间意味着更强的数据吞吐量以及更好的散热表现。

此外,台积电还展示了一系列高性能集成解决方案,其中包括专为HBM4设计的N12和N3制程基础裸片(Base Die),以及运用紧凑型通用光子引擎技术的硅光子学整合方案。这些技术将进一步推动高性能计算和人工智能领域的发展。

黄茹 供稿自

<https://www.bntimes.com/companies/tsmc-unveils-a14-process-and-system-on-wafer-x-a-leap-forward-for-ai-and-high-performance-computing>

原文标题: TSMC Unveils A14 Process and System on Wafer-X: A Leap Forward for AI and High-Performance Computing

产业动态

[人工智能]NVIDIA 计划开发 AI 超级计算机

2025年4月14日，英伟达宣布将在美国本土投资超过5000亿美元，并在未来四年内携手台积电等多家巨头，打造人工智能基础设施，以推进AI芯片和超级计算机的本土化制造，强化供应链安全。

作为新型数据中心的核心，英伟达AI超级计算机专为处理人工智能而设计，是推动AI产业快速发展的关键基础设施。未来几年，英伟达计划建成数十座“千兆瓦AI工厂”，用于制造AI芯片和超级计算机，预计将创造数十万个就业岗位，为美国经济注入数万亿美元的安全保障。

未来四年内，英伟达计划与台积电、富士康、纬创、安靠和日月光半导体深度合作，共同推进AI芯片与超级计算机的本土化制造。目前，英伟达Blackwell芯片已经开始在台积电亚利桑那州凤凰城工厂投入生产。此外，英伟达还将与富士康、纬创在休斯敦和达拉斯建设超级计算机制造工厂，这两座工厂预计将在未来12-15个月内逐步启动大规模量产。同时，英伟达将与安靠和日月光半导体合作在亚利桑那州开展封装测试业务。

黄茹 供稿自

<https://www.nextgov.com/emerging-tech/2025/04/nvidia-announces-plans-fully-us-manufactured-ai-supercomputers/404541/>

原文标题：NVIDIA announces plans for fully US-manufactured AI supercomputers

[人工智能]OpenAI 宣布开发开源 AI 模型

2025年4月1日，OpenAI宣布正在构建一款开放权重（Open-Weight）的生成式AI模型，将具备与o3-mini相当的推理能力，以应对在开源领域来自DeepSeek和Meta等企业的竞争压力。

OpenAI曾是闭源模型的捍卫者，认为开源模型风险太大，容易被不良行为者或者国外竞争对手恶意操纵。自2019年OpenAI成立营利性子公司以来，一直走的是闭源路线，通过ChatGPT订阅费用以及付费API业务来赚取利润。

事实证明，许多企业和政府不愿意在其无法控制的模型上构建他们的AI产品或服务，这对坚持闭源路线的OpenAI构成了巨大压力。开源则能为OpenAI打开更多的市场，吸引更多开发者和企业用户，解决闭源模型带来的用户流失问题。

唐衡 供稿自

<https://techxplora.com/news/2025-04-shift-openai-ai.html>

原文标题：In shift, OpenAI announces open AI model

[RISC-V]香山高性能处理器设计融入主流编译器基础设施

2025年4月22日，LLVM开源项目接受了香山开源RISC-V处理器项目提交的“昆明湖”核心硬件描述代码（Commit 0e3e0bf），标志着“昆明湖”高性能RISC-V处理器核获得通用编译器项目正式支持。

全球开发者在LLVM框架下可直接编译和优化针对这款对标ARM Neoverse N2的高性能处理器核心的应用程序。“昆明湖”作为香山项目的第三代核心，采用7nm先进工艺，主频高达3GHz，SPECint2006评分达15分/GHz。此次提交点亮了国际RISC-V编译生态主线，实现开源高性能RISC-V处理器与国际生态平滑高效接轨。目前，已有企业基于香山核心开发数据中心AI芯片、RISC-V服务器集群，合并后的工具链能有效提升基于香山“昆明湖”的RISC-V软件开发效率。

此外，香山项目GitHub仓库已获3200+星标，此次合并开启了开源处理器与开源编译器的生态联动，形成跨领域融合新篇章。

黄茹 供稿自

https://mp.weixin.qq.com/s/jPcWA_69wRqjbM4EZtvhJg

原文标题：“香山昆明湖”硬件描述并入LLVM主线，香山高性能处理器设计正式融入全球主流编译器基础设施

版权及合理使用声明

《科学研究动态监测快报》（以下简称《监测快报》）是由中国科学院文献情报中心、中国科学院成都文献情报中心、中国科学院武汉文献情报中心以及中国科学院兰州文献情报中心和中国科学院上海生命科学信息中心按照主要科学研究领域分工编辑的科学研究进展动态监测报道类信息快报。

《监测快报》遵守国家知识产权法的规定，保护知识产权，保障著作权人的合法权益，并要求参阅人员及研究人员遵守中国版权法的有关规定，严禁将《监测快报》用于任何商业或其他营利性用途。读者在个人学习、研究目的中使用信息报道稿件，应注明版权信息和信息来源。未经编辑单位允许，有关单位和用户不能以任何方式全辑转载、链接或发布相关科学领域专辑《监测快报》内容。有关用户单位要链接、整期发布或转载相关学科领域专辑《监测快报》内容，应向具体编辑单位发送正式的需求函，说明其用途，征得同意，并与具体编辑单位签订服务协议。

欢迎对《科学研究动态监测快报》提出意见与建议。

信息科技专辑：

编辑出版：中国科学院成都文献情报中心

联系地址：四川省成都市天府新区群贤南街 289 号（610299）

电话：（028）85235556

电子邮件：casit@clas.ac.cn

内部资料

中国科学院成都文献情报中心

新一代信息科技战略研究中心

电话：028-85235075

E-mail: casit@clas.ac.cn

地址：四川省成都市群贤南街289号, 610299