



人工智能发展月报

2023年4月

总第 30 期

洞察前沿 发现情报

报告说明

本报告依托科技情报大数据挖掘与服务系统平台 AMiner、新闻事件分析挖掘和搜索系统 NewsMiner，以及人工智能主流新闻网站及公众号，从 AI 学术会议、重大科研进展、人物动态、最新报告发布等角度，分析挖掘了每月人工智能领域所发生的、对 AI 领域技术发展产生重大推动作用的事件，旨在帮助 AI 领域研究人员和业内人士及时追踪最新科研动态、洞察前沿热点。

本期导读

4 月份，人工智能业界事件仍然以大模型为焦点，并呈现出两股强大力量：一是大模型研发进展与产品应用令人应接不暇，二是全球对 AI 发展的政策支持与监管并行加强。本月的整体趋势如下图所示。其中，18 日热度最高，主要由于除了最高检发文加大对 AI 核心技术的司法保护力度、Meta 官宣 DINOv2 等之外，引入大模型的 AI 办公助手于当日扎堆密集发布，例如石墨、腾讯文档、印象笔记、钉钉、百度如流、北大 Excel、WPS AI 等。此外，Google DeepMind 合并成立、复旦 MOSS 以及清华 BioMedGPT 模型开源等事件也受到较高关注。

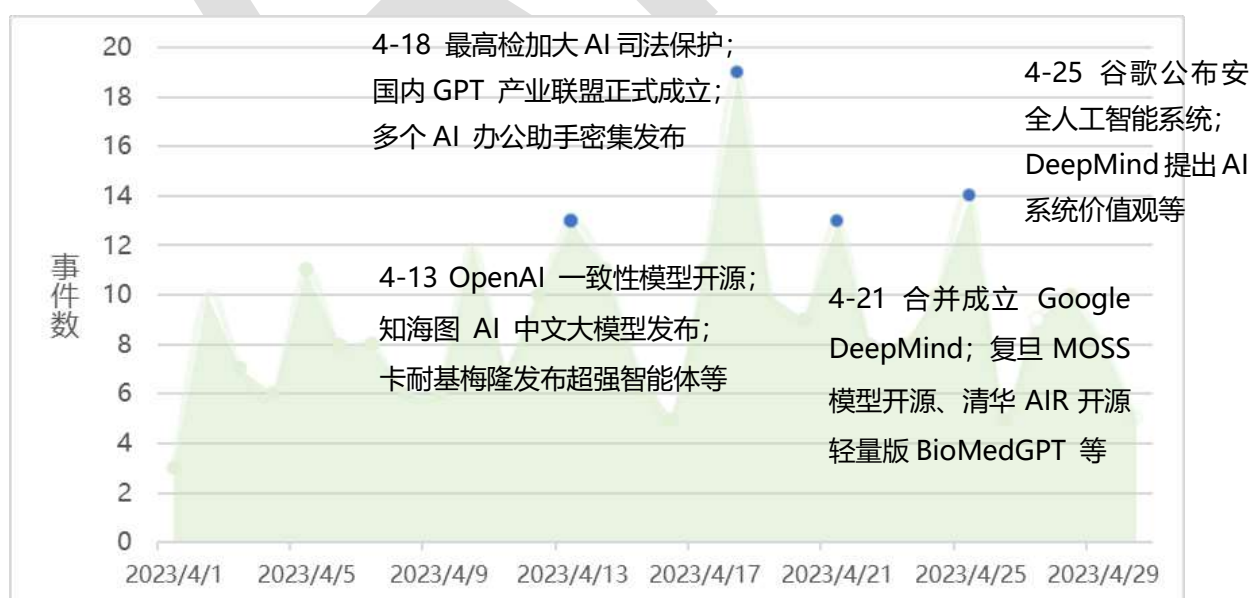


图 1 人工智能领域 2023 年 4 月新闻事件趋势

本月值得关注的 AI 科研进展

机器学习

- 牛津大学开源框架 PureJaxRL 实现超 4000 倍的加速
- 斯坦福大学提出在高层级上增加序列长度的新方法
- 港科大 LMFlow 团队开源高效对齐算法 RAFT [【更多】](#)

计算机视觉

- META 发布视觉模型 Segment Anything Model (SAM) 、DINOv2 分割检索模型、开放类分割模型 OVSeg 等
- 谷歌提出一种能够高效且稳定训练大规模 ViT 模型的方法、扩散模型推理加速新方法等
- 智源推出通用分割模型 SegGPT [【更多】](#)

AIGC

- UC 伯克利发布对话模型 Koala
- Hugging Face 发布 300 亿参数的 HuggingChat
- 微软亚洲研究院发布超长视频生成模型 NUWA-XL
- 阿里云大模型“通义千问”开启企业邀测
- 百度、钉钉、字节等多家发布 AI 办公助手 [【更多】](#)

智能芯片

- 谷歌发布定制芯片 TPU、WebGPU
- 微软正在开发人工智能芯片雅典娜
- 火山引擎自研 DPU 已部署上万台 DPU 服务器 [【更多】](#)

智能机器人

- MIT 发明的机器手可通过感知形状来识别东西
- 波士顿动力 Spot 机械狗成功集成 ChatGPT
- 阿里云正在实验将千问大模型接入工业机器人 [【更多】](#)

AI 应用

- 利用 AI 算法实现精神分裂症的个体精准化分型
- 蛋白质领域首次使用对比学习准确预测酶功能
- 华盛顿大学开发了一款新型蛋白质设计软件
- 清华团队为“AI 司机”量身“设计考题”
- 沃尔玛用 AI 聊天机器人和供应商讨价还价 [【更多】](#)

具体详情，请参见下文

I 目录

本期导读.....	2
一、AI 顶会动态	7
1. 顶会 IJCAI 2023 放榜，论文接收率 15%	7
2. 顶会 ICML 2023 放榜，论文录用率 27.9%	7
二、AI 重要政策法规	7
【国内】	7
1. 政治局会议首提“通用人工智能” 明确长期产业趋势	7
2. 科技部重组后首发力：成立 AI 推进办，启动 AI 重大项目	8
3. 国家基金委：下一代人工智能方法重大研究计划 2023 年度项目指南	8
4. 工信部：将研究制定系统深入推进智能制造的行动计划	9
5. 网信办：《生成式人工智能服务管理办法（征求意见稿）》	10
6. 最高检：加大对集成电路、AI 等重点领域核心技术司法保护力度 ..	10
7. 国内 GPT 产业联盟正式成立	11
8. 上海对人工智能关键技术项目最高支持两千万	12
9. 北京经济和信息化局发文促进机器人创新发展	12
【国外】	13
10. 英国政府出台政策文件并拨款 1 亿 支持人工智能监管创新	13
11. 美国国家电信和信息管理局拟制定人工智能问责机制	13
12. 意大利禁用 ChatGPT 后：欧盟各国跟进效仿	14
13. 欧洲议会拟出台法案严格管理人工智能	15
14. 欧盟达成《人工智能法》协议，为生成式 AI 设立版权规则	15
15. 日本计划建立人工智能国家战略委员会	15
三、AI 最新报告发布	16
1. 斯坦福 2023 AI Index：大模型作者一半来自美国	16
2. 国家网信办：我国数字经济规模稳居世界第二	17
3. 阿里研究院、智谱 AI：2023 全球数字科技十大趋势	18
4. 中金研究院：中国人工智能创新能力面临三大挑战	18

四、榜单奖项	19
1. 全球 TOP 1000 计算机科学家公布：中国 96 位学者上榜	19
2. 2023 Meta 博士奖研金名单出炉：华人学者超 1/3	21
3. 亚马逊研究奖名单公布：华人获奖者近 1/3	22
4. 福布斯：最有前途的人工智能公司 AI 50 榜单	23
五、国内高校动态	24
1. 教育部：全国 498 所高校成功申报人工智能专业	24
2. 教育部：全国 71 所高校成功申报数字经济专业	25
3. 高校动态	25
六、AI 人才动态	27
【学术界】	27
1. 微软黄学东入选美国艺术与科学院新增院士	27
2. 陈怡然荣获 2022 Charles A. Desoer 技术成就奖	28
【产业界】	29
3. Jeff Dean 将担任 Google DeepMind 的首席科学家	29
4. 搜狗前 CEO 王小川官宣大模型创业，对标产品今年推出	29
5. IDEA 研究院张家兴入局大模型创业	30
6. 字节技术大牛创业打造聚焦 AIGC 的视觉多模态算法平台	30
7. 小米任命栾剑为 AI 实验室大模型团队负责人	31
七、子领域重要科研进展	31
1. 机器学习	31
2. 计算机视觉	36
3. 自然语言处理	42
4. AIGC/生成式 AI	45
(1) 基础模型	45
(2) 垂直领域基础模型	51
(3) 产品应用	54
5. 智能芯片	60

6. 智能机器人.....	65
7. 语音识别.....	67
8. 信息检索与推荐	67
9. 安全与伦理.....	68
10. AI 应用	69
(1) AI+ 医疗	69
(2) AI+ 自然科学	69
(3) 元宇宙/数字人	71
(4) AI+ 金融	71
(5) AI+ 社会生活娱乐.....	72



一、AI 顶会动态

1. 顶会 IJCAI 2023 放榜，论文接收率 15% ¹

4 月 21 日，顶会 IJCAI 2023 公布本届论文录用情况，共收到 4566 份提交全文，接收率大约 15%。IJCAI 2023 将于 8 月 19 日至 25 日在澳门召开。

2. 顶会 ICML 2023 放榜，论文录用率 27.9% ²

4 月 25 日，ICML 2023 公布了今年录取情况，共有 6538 份论文提交，1827 篇论文被接收，录用率 27.9%。作为对比，去年共收到 5630 投稿，接收 1117 篇 short oral，118 篇 long oral，录用率为 21.94%。第 40 届机器学习顶会 ICML 2023 将在 7 月 23-29 在夏威夷召开。



二、AI 重要政策法规

【国内】

1. 政治局会议首提“通用人工智能” 明确长期产业趋势 ³

中共中央政治局 4 月 28 日召开会议，分析研究当前经济形势和经济工作。

¹ [IJCAI 2023 放榜！接收率 15%：中稿全凭运气？ rebuttal 没用？ 审稿人只盯负面评价？ 4 月 21 日](#)

² [录用 1827 篇，接收率 27.9%，ICML 2023 接收结果公布，4 月 25 日](#)

³ [中央政治局会议：重视通用人工智能发展，营造创新生态，重视防范风险|人工智能 新浪财经](#)；[政治局会议首提“通用人工智能” 明确长期产业趋势](#)，4 月 28 日

中共中央总书记习近平主持会议。会议指出，要夯实科技自立自强根基，培育壮大新动能；要重视通用人工智能发展，营造创新生态，重视防范风险。

2. 科技部重组后首发力：成立 AI 推进办，启动 AI 重大项目⁴

日前，科技部高新技术司司长陈家昌在国务院新闻办公室新闻发布会上表示，在人工智能方面，科技部加强顶层设计，成立人工智能规划推进办公室，启动实施新一代人工智能重大科技项目，在数字孪生、数字制造、智慧医疗等方面都做了相应部署。同时，针对人工智能发展过程中一些风险和问题，制定发布《新一代人工智能治理原则——发展负责任的人工智能》和《新一代人工智能伦理规范》，推动科技向善、造福人类。下一步，科技部将充分发挥新一代人工智能规划推进办公室的协调作用，整合项目、平台、人才等资源，形成推进的政策合力。

3. 国家基金委：下一代人工智能方法重大研究计划 2023 年度项目指南⁵

日前，国家自然科学基金委发布《可解释、可通用的下一代人工智能方法重大研究计划 2023 年度项目指南》。根据《指南》，该重大研究计划聚焦的核心科学问题为：深度学习的基本原理；可解释、可通用的下一代人工智能方法；面

⁴ [科技部重组后首发力：成立 AI 推进办，启动 AI 重大项目|人工智能 网易](#)，4 月 10 日；[科技部：已启动实施新一代人工智能重大科技项目—科学网](#)，4 月 3 日

⁵ [最高 300 万元/项！国家基金委发布下一代人工智能方法重大研究计划 2023 年度项目指南-高校科技-中国教育和科研计算机网](#)，4 月 4 日；[基金委发布下一代人工智能方法重大研究计划项目指南—科学网](#)

向科学领域的下一代人工智能方法的应用。该重大研究计划 2023 年度拟资助培育项目 25 ~ 30 项，直接费用资助强度约为 80 万元/项，资助期限为 3 年；拟资助重点支持项目 6 ~ 8 项，直接费用资助强度约为 300 万元/项，资助期限为 4 年。《指南》明确，围绕核心科学问题，以总体科学目标为牵引，拟以培育项目的方式资助探索性强、选题新颖的申请项目，涵盖“深度学习的表示理论和泛化理论”、“深度学习的训练方法”、“微分方程与机器学习”、“隐私保护的机器学习方法”、“图神经网络的新方法”等十大研究方向；拟以重点支持项目的方式资助前期研究成果积累较好、对总体科学目标在理论和关键技术能发挥推动作用、具备产学研用基础的申请项目，涵盖“经典数值方法与人工智能融合的微分方程数值方法”、“复杂离散优化的人工智能求解器”、“开放环境下多智能体协作的智能感知理论与方法”等八大研究方向。

4. 工信部：将研究制定系统深入推进智能制造的行动计划⁶

4 月 7 日，2023 年国家智能制造标准化总体组、专家咨询组全体会议在北京召开。工业和信息化部党组成员、副部长辛国斌表示，工业和信息化部将研究制定系统深入推进智能制造的行动计划，深化智能制造标准领航行动，希望总体组、专家咨询组继续发挥好统筹协调和战略指导作用。坚持问题导向，加快重点

⁶ [工信部：将研究制定系统深入推进智能制造的行动计划 央视网](#)，4 月 7 日

标准研制和行业标准应用，打造一批基于典型应用场景的标准化解决方案。坚持创新驱动，面向前沿技术领域，加快布局一批引领性国家标准和行业标准。坚持统筹协调，开展标准应用试点和贯标工作，开发应用智能制造标准群和工具箱，完善智能工厂、智慧供应链、智能制造先行区等遴选标准。坚持开放合作，积极参与国际标准制修订，在国际标准化组织中贡献更多中国方案。

5. 网信办：《生成式人工智能服务管理办法（征求意见稿）》⁷

4月11日，为促进生成式人工智能技术健康发展和规范应用，根据《中华人民共和国网络安全法》等法律法规，国家互联网信息办公室起草了《生成式人工智能服务管理办法（征求意见稿）》，现向社会公开征求意见。意见反馈截止时间为2023年5月10日。

6. 最高检：加大对集成电路、AI等重点领域核心技术司法保护力度⁸

4月18日，最高人民检察院印发《关于加强新时代检察机关网络法治工作的意见》。其中，围绕聚焦数字经济健康发展，依法保护和规范数字技术、数字产业和数字市场方面，《意见》指出，要发挥知识产权检察办公室综合履职的优

⁷ [国家互联网信息办公室关于《生成式人工智能服务管理办法（征求意见稿）》公开征求意见的通知，4月11日](#)

⁸ [最高检：加大对集成电路、人工智能等重点领域核心技术司法保护力度，4月18日；最高检印发《关于加强新时代检察机关网络法治工作的意见》 中华人民共和国最高人民检察院](#)

势，突出加大对集成电路、人工智能、大数据、云计算等重点领域核心技术的司法保护力度，依法加强对计算机软件、数据库、网络域名、数字版权、数字内容作品等网络知识产权的司法保护，依法惩治滥用知识产权排除市场竞争、阻碍创新的行为，推动数字技术成果的创新发展和转化。《意见》指出，要加强对“元宇宙”、人工智能、区块链、云计算等新技术新业态相关法律问题的前瞻研究，强化安全风险综合研判，准确把握创新发展与违法犯罪的界限，及时发现、精准惩治以新技术新业态为幌子实施的各类犯罪活动，既保护创新发展，又有效防范各类风险叠加演变。

7. 国内 GPT 产业联盟正式成立⁹

4 月 18 日，在 2023 全球元宇宙大会上，GPT 产业联盟正式宣告成立。GPT 产业联盟由中国移动通信联合会、中国电信、中国移动、中国联通、中国广电等单位共同发起成立。联盟目标：促进全球范围内的 AI 技术交流与合作；制定 AI 领域的共同标准和最佳实践；鼓励公平、开放和透明的 AI 技术研究和应用；积极应对 AI 技术带来的安全、伦理和社会挑战。

⁹ [GPT 产业联盟正式成立：含中国电信、移动、联通、广电等企业|中国电信|产业联盟|广电 新浪新闻](#)，4 月 18 日

8. 上海对人工智能关键技术项目最高支持两千万¹⁰

近日，上海市人民政府办公厅印发《关于新时期强化投资促进加快建设现代化产业体系的政策措施》。上面显示，围绕芯片设计、制造、封测、装备、材料等领域，积极招引硬实力优质企业落地；对引进符合条件的零部件、原材料等集成电路装备材料重大项目和 EDA 等高端软件项目，给予不超过项目投资的 30%、最高 1 亿元的支持。其中，瞄准人工智能领域具有全球影响力的重大原创成果、前沿理论、龙头企业等，加快招引智能芯片、核心算法、操作系统和基础软件等重点项目落地。对引进符合条件的人工智能关键技术项目，给予不超过项目投资的 30%、最高 2000 万元的支持。

9. 北京经济和信息化局发文促进机器人创新发展¹¹

近日，北京市经济和信息化局印发《北京市机器人产业创新发展行动方案（2023-2025 年）（征求意见稿）》。根据《方案》，到 2025 年，北京市培育 100 种高技术高附加值先进产品、100 种具有全国推广价值的示范场景、100 家专精特新“小巨人”企业，建成 5 个国家级机器人产业公共服务平台，全市机器人核心产业收入达到 300 亿元以上，打造国内领先、具有国际先进水平的机器人产业集群。根据《方案》，将组建北京市人形机器人产业创新中心。

¹⁰ [泰酷辣！上海对人工智能关键技术项目最高支持两千万|人工智能|上海市|元宇宙 新浪科技，4 月 25 日](#)

¹¹ [北京发文促进机器人创新发展，2025 年核心产业收入超 300 亿元，4 月 26 日](#)

【国外】

10. 英国政府出台政策文件并拨款 1 亿 支持人工智能监管创新¹²

3 月 29 日，英国政府发布《支持创新的人工智能监管方式》（A pro-innovation approach to AI regulation）政策文件，旨在寻求建立社会共识，加深公众对尖端技术的信任，使企业更好地创新、发展，从而创造更多就业机会。文件概述了人工智能监管应考虑的五项原则，即安全性和稳健性；透明度和可解释性；公平性；问责制和管理；可竞争性和补救性。文件指出，人工智能的使用应符合英国现有的法律，不得歧视个人或产生不公平的商业结果；需要采取措施确保对人工智能的使用方式进行适当的监督，并对结果进行明确的问责。

4 月 24 日，英国政府正式宣布将初步拨款 1 亿英镑（约 8.63 亿人民币）加速推动人工智能（AI）发展和开发。一个由英国专家组成的工作组将协助创建基础模型，将人工智能推广至医疗、教育等领域。

11. 美国国家电信和信息管理局拟制定人工智能问责机制¹³

4 月 12 日，美国政府表示，正在就人工智能（AI）系统的潜在问责措施征

¹² [英国政府出台政策文件 支持人工智能监管创新 - 西部网（陕西新闻网）](#)，4 月 13 日；[英国政府宣布拨款 1 亿英镑推动人工智能发展 开发 模型 领域](#)，4 月 24 日

¹³ [美国拟制定人工智能问责机制 多国考虑将 ChatGPT 纳入监管](#)，第一财经，4 月 12 日

求公众意见。美国国家电信和信息管理局（NTIA）称，希望了解是否可以采取措施来保证人工智能系统是“合法、有效、合乎道德且安全的，并且在其他方面值得信赖”。NTIA 是商务部的一个机构，就电信和信息政策向白宫提供建议。

12. 意大利禁用 ChatGPT 后：欧盟各国跟进效仿¹⁴

意大利禁用 ChatGPT 的做法，引发了欧洲其他国家的关注，近日，法国、德国、爱尔兰等国家都已开始准备效仿意大利的做法，加强对 ChatGPT 的监管。4 月 3 日，法国和爱尔兰的隐私监管机构已经与意大利的同行取得联系，了解更多有关 ChatGPT 禁令的依据。与此同时，德国联邦数据保护专员乌尔里希·凯尔伯（Ulrich Kelber）表示，德国可能会效仿意大利的做法，出于对数据安全保护的考量而选择禁止 ChatGPT。但瑞典的隐私监管机构表示，目前还没有禁止 ChatGPT 的计划，也没有与意大利的监管机构取得过联系。西班牙的监管机构称，并不排除未来展开调查的可能性。意大利个人数据保护局（DPA）曾宣布从 3 月 31 日起禁止使用聊天机器人 ChatGPT，限制 OpenAI 处理意大利用户信息数据，同时对其隐私安全问题立案调查。

¹⁴ [意大利禁用 ChatGPT 后：欧盟各国跟进效仿，本国副总理批“监管过度” - 智东西](#)，4 月 4 日

13. 欧洲议会拟出台法案严格管理人工智能¹⁵

近日，在欧盟逐步制定全世界最具约束性的人工智能管理方法之际，欧洲议会准备对人工智能的使用实施严厉的新措施，其中包括要求聊天机器人开发商披露他们是否使用了受版权保护的材料。此举旨在让内容创作者有权获得报酬。欧洲议会议员提出的一项有争议的提案是，禁止公共场所任何情况下使用面部识别技术。

14. 欧盟达成《人工智能法》协议，为生成式 AI 设立版权规则¹⁶

欧洲议会成员已经弥合分歧，就《人工智能法（The AI Act）》提案达成临时政治协议，要求部署 ChatGPT 等生成式人工智能工具的公司披露用于开发其系统的受版权保护的材料。该协议为可能是世界上第一部综合性人工智能法律的《人工智能法》出台铺平道路。在定于 5 月 11 日进行的关键委员会投票前，提案文本可能仍需在技术层面上进行细微调整，6 月中旬预计将进行全体投票。

15. 日本计划建立人工智能国家战略委员会¹⁷

日本政府决定设立新的“战略会议”，负责讨论与人工智能（AI）相关的国家战略。该委员会将发挥指挥塔作用，为有关人工智能的整体政策提供基本指导，

¹⁵ [欧洲议会拟出台法案严管 AI 人工智能 技术 的材料，4 月 16 日](#)

¹⁶ [欧盟达成《人工智能法》协议，为生成式 AI 设立版权规则 未来 2% 澎湃新闻，4 月 28 日](#)

¹⁷ [日本计划建立人工智能国家战略委员会 腾讯新闻，4 月 28 日](#)

包括 ChatGPT。预计该委员会还将从人工智能的利用和促进研发以及收紧监管的角度，讨论涉及人工智能的挑战。该委员会还将讨论在教育、行政和私营企业中使用对话式人工智能所涉及的挑战，以及如何促进日本企业和机构的人工智能研发。



三、AI 最新报告发布

1. 斯坦福 2023 AI Index：大模型作者一半来自美国¹⁸

近日，斯坦福大学以人为本人工智能研究所（Stanford HAI）发布了《[2023 年人工智能指数报告](#)》（Artificial Intelligence Index Report 2023）。这是该机构发布的第 6 份年度报告。AI Index 表明，在过去一年里，大量 AI 工具已经成为主流，从 ChatGPT 等聊天机器人到 Midjourney 等图像生成软件。但是关于如何部署这项技术以及如何平衡风险和机会的决定权牢牢掌握在企业参与者手中。AI Index 还指出，多年来，学术界在开发最先进的 AI 系统方面处于领先地位，但现在工业界已经牢牢接管。全球 AI 出版物的数量从 2010 年到 2021 年，总数量翻了一番多。自 2015 年以来，模式识别方面的论文数量大约翻了一番，而机器学习论文的数量大约翻了两番。过去 12 年中中美合作次数最

¹⁸ [斯坦福 2023 AI Index 回顾十年进展：大模型作者一半来自美国机构、中国期刊论文领先 澎湃](#)
[号·湃客 澎湃新闻-The Paper](#)，4 月 5 日；[2023 年斯坦福 AI Index 公布：细数 AI 最新进展及年度发展趋势](#)
[学术头条](#)

多，自 2010 年以来增长了大约四倍；然而，从 2020 年到 2021 年，中美合作总数仅增长了 2.1%。AI 期刊论文的发表数量，中国始终保持领先地位，2021 年为 39.8%，其次是欧盟和英国（15.1%）、美国（10.0%）。大型语言模型和多模态模型的研究人员主要来自美国机构（54.2%）。

2. 国家网信办：我国数字经济规模稳居世界第二¹⁹

4 月 27 日，第六届数字中国建设峰会开幕式上，国家互联网信息办公室发布了《数字中国发展报告(2022 年)》。报告指出，目前数字经济成为中国稳增长促转型的重要引擎。2022 年，中国数字经济规模达 50.2 万亿元(人民币)，稳居世界第二，同比名义增长 10.3%，占 GDP 比重提升至 41.5%。截至 2022 年底，中国开通 5G 基站 231.2 万个，5G 用户 5.61 亿户，全球占比超 60%；移动物联网终端用户达 18.45 亿户，成为全球主要经济体中首个实现“物超人”的国家；算力总规模达到 180EFLOPS(每秒 18000 京次浮点运算)，位居世界第二，年增长率近 30%。中国 2022 年数据产量达 8.1ZB，同比增长 22.7%，全球占比 10.5%，位居世界第二。浙江、北京、广东、江苏、上海、福建、山东、天津、重庆、湖北等地区数字化综合发展水平位居全国前 10 名。

¹⁹ 《数字中国发展报告（2022 年）》发布：数字领域国际合作凝聚广泛共识-中新网，4 月 27 日

3. 阿里研究院、智谱 AI：2023 全球数字科技十大趋势²⁰

阿里研究院、智谱 AI 近日联合发布《2023 全球数字科技技术发展研究报告》。报告基于 AMiner 科技情报平台的数据，利用文献计量方法，为数字科技研究前沿“画像”，通过对 2012 年-2021 年数字科技领域 TOP1%核心论文的持续跟踪，筛选总结了 2023 全球数字科技十大趋势：生物大数据、生成式对抗网络算法（GANs）、沉浸式扩展现实（Extended Reality, XR）、量子计算机、AI 破解蛋白质和基因结构、移动边缘网络、可解释的 AI、联邦学习、混合计算和能源区块链技术。报告还发现，从全球数字技术论文的“量与质”方面来看，中美数字科技论文整体影响力大致相当。尽管中国数字科技领域论文数量（506,775 篇）与美国（525,794 篇）还存在 4%的差距，但是平均被引量中国比美国多 4 次/篇，这说明中国数字科技领域论文整体影响力与美国大致相当。

4. 中金研究院：中国人工智能创新能力面临三大挑战²¹

近日，中金研究院发布《CGI 深度 | 竞争人工智能：从论文突破性看中国人工智能创新水平》。该报告发现，总量上，中国人工智能的论文发表量在 2017 年反超美国成为全球第一，突破性论文发表量在 2016 年升至全球第二、2020 年

²⁰ [重磅发布|《2023 全球数字科技技术发展研究报告》：2023 全球数字科技十大趋势有哪些？](#) 搜狐，4 月 21 日

²¹ [竞争人工智能：从论文突破性看中国人工智能创新水平 - 华尔街见闻](#)，4 月 4 日；[CGI 深度 | 竞争人工智能：从论文突破性看中国人工智能创新水平](#)

国际份额达 17%；结构上，中国在模式识别、计算机视觉、数据挖掘、语音识别领域的突破性表现尚可，而在算法、自然语言处理和人机交互领域相对薄弱。国际对比来看，中美突破性创新的差距逐步缩小、但仍然较大。大国竞争下，中美人工智能论文的合作比例在 2018 年后出现下滑，但并未对我国突破性创新造成很大的消极影响，与美国的突破性论文合作份额被其他国家取代。报告还指出，过去二十年，我国在人工智能领域的突破性创新能力取得了长足的进步，已跃居全球第二的水平，但在人工智能的创新能力方面与美国差距仍然较大，并且面临三大挑战：第一，科研机构论文发表量多，但突破性成果少；第二，科研人员的规模大，但尖端人才少，部分前沿短板领域科研人员相对不足；第三，商业机构在人工智能突破性创新中的缺位问题突出，限制了突破性人工智能产品的推出。



四、榜单奖项

1. 全球 TOP 1000 计算机科学家公布：中国 96 位学者上榜²²

近日，Research (Guide2Research) 发布了 2023 年世界顶尖 1000 名计算机科学家排名。排名基于包括 OpenAlex 和 CrossRef 等各种数据源整合的数据。美国 583 名科学家实力霸榜，占 58.3%；中国和英国位于第 2 名和第 3 名，

²² [重磅！全球 TOP 1000 计算机科学家公布：中国 96 位学者上榜！港理工张磊居大陆科学家之首](#)，新智元，4 月 1 日

分别有 96 名和 64 名科学家，占比为 9.2%和 6.4%。在排名前 1%的科学家中，有 5 名研究人员为美国机构工作，2 名来自加拿大大学，还有来自德国、英国和澳大利亚的学者各 1 名。斯坦福大学的顶尖计算机科学家数量跃居榜首，共有 38 名。中国科学家 96 人上榜（大陆 92 人+台湾 4 人），港理工张磊居大陆科学家之首。美国伊利诺伊大学香槟分校计算机系教授韩家炜 (Jiawei Han) 是唯一一位进入全球 TOP5 的华人计算机科学家。华人学者在世界 TOP20 中占 3 席，分别是韩家炜（美国伊利诺伊大学香槟分校，数据挖掘方向）、俞士纶（伊利诺大学芝加哥分校，数据挖掘与管理方向）、黄煦涛（伊利诺伊大学厄巴纳香槟分校，信息和信号处理方向）。本次上榜的大陆地区科学家排名前 10 如下。

国内排名（国际排名）	学者	机构
1 (30)	张磊	香港理工大学
2 (32)	汤晓鸥	商汤
3 (39)	王晓刚	香港中文大学
4 (40)	徐泽水	四川大学
5 (43)	杨易	浙江大学
6 (57)	张宏江	源码资本
7 (60)	杨强	香港科技大学
8 (70)	张大鹏	香港中文大学（深圳）
9 (78)	李学龙	西北工业大学
10 (89)	沈定刚	上海科技大学

2. 2023 Meta 博士奖研金名单出炉：华人学者超 1/3²³

4月5日，Meta 公布了 2023 年博士奖研金（PhD Fellowship）名单。获奖者将获得两个学年的全额学费和其他费用以及 42,000 美元的助学津贴，以及与 Meta 研究人员接触交流机会。今年，组委会收到了来自全球 100 多所大学的 3,200 多份申请，并从 12 所大学中挑选出了 21 名获奖者，其中超过 1/3 是华人博士生。本次获奖的 8 位华人博士生名单如下。

博士姓名	所在机构	研究方向
Mark Zhao	斯坦福大学	协同设计计算机系统和架构、数据存储和摄取系统
Victoria Lin	卡内基梅隆大学	因果推断方法
Yiling Qiao	马里兰大学帕克分校	基于物理的模拟，可微分模拟及其在虚拟现实 / 增强现实、图形学、机器人学等领域的应用
Jinkun Cao	卡内基梅隆大学	对视频中人类运动的分析、建模和合成
Yue Jiang	芬兰阿尔托大学和芬兰人工智能中心	将计算方法应用于自适应用户界面
Shuaifeng Jiang	亚利桑那州立大学	无线通信和传感、毫米波和大规模 MIMO 系统、机器学习
Dawei Liang	得克萨斯大学奥斯汀分校	可穿戴计算、音频识别和以人为中心 AI 的交叉领域
Xuechen Li	斯坦福大学	改进机器学习系统的隐私保护机器学习和安全计算

²³ [2023 Meta 博士奖研金名单出炉：华人学者超 1/3](#)，机器之心，4月6日

3. 亚马逊研究奖名单公布：华人获奖者近 1/3 ²⁴

近日，亚马逊研究奖（Amazon Research Awards, ARA）获奖名单公布，今年共有 79 位获得者，来自 14 个国家 54 所大学。获奖者可以获得不受限制的资金，还可以公开访问 300 多个亚马逊公共数据集，并可以使用亚马逊的 AI/ML 服务和工具。获奖者还将与亚马逊专家建立联系，以获得咨询和建议，还可以参加亚马逊举办的活动、培训课程等。

获奖者姓名	所在机构	研究课题
Yan Huang	卡内基梅隆大学	对抗从人类决策中继承的算法偏差
Zhihao Jia	卡内基梅隆大学	利用系统优化方法构建可持续 AI
李磊	加州大学圣芭芭拉分校	少量样本情况下实时稳健的同声传译
Xiaoyi Lu	加州大学默塞德分校	为分布式深度学习训练扩展集体通信
Yunan Luo	佐治亚理工学院	用于稳健药物筛选的经校准的、可解释的几何深度学习
Qing Qu	密歇根大学安娜堡分校	通过神经崩溃（neural collapse）进行深度表征学习
宋晓东 Dawn Song	加州大学伯克利分校	FedOps: 可信赖的联邦学习的抽象
Xiaolong Wang	加州大学圣地亚哥分校	加州大学圣地亚哥分校
Ce Zhang	苏黎世联邦理工学院	FedOps: 可信赖的联邦学习抽象
Tianyi Zhang	普渡大学	Human-in-the-loop 中的深度学习优化，以达

²⁴ [亚马逊研究奖获奖名单公布：李磊、朱俊彦等多位华人在列，4月19日](#)

		到更好的可用性，透明度和用户信任
Yiying Zhang	加州大学圣地亚哥分校	用 zero 激活训练深度神经网络
Jishen Zhao	加州大学圣地亚哥分校	基于语义的大型语言模型文档结构识别
Ben Zhao	芝加哥大学	深度神经网络的数字取证
Heather Zheng	芝加哥大学	深度神经网络的数字取证
朱俊彦	卡内基梅隆大学	大规模生成模型的组合个性化
Jia Zou	亚利桑那州立大学	加速机器学习推理查询的编译框架
Sicun Gao	加州大学圣地亚哥分校	蒙特卡罗树决策方法
Limin Jia	卡内基梅隆大学	通过程序合成实现一行 rust 验证
Xiaorui Liu	北卡罗来纳州立大学	基于页面显示的推荐和广告的深度强化学习混合排序
汤继良	密歇根州立大学	基于页面显示的推荐和广告的深度强化学习混合排序
Fan Zhang	布里斯托大学	扩展参数空间中的通用深度视频质量评估

4. 福布斯：最有前途的人工智能公司 AI 50 榜单 ²⁵

近日，《福布斯》与红杉资本 (Sequoia) 和 Meritech Capital 合作评选发布了最有前途的利用人工智能开展业务的未上市公司榜单。2023 年总共收到了

²⁵ [福布斯发布 AI 50 榜单：最有前途的人工智能公司，4 月 14 日](#)

近 800 份申请，评选流程包括由一个定量算法和一个定性评审团来评判他们的商业前景和对人工智能的技术应用。上榜的 50 家公司总共获得了 272 亿美元的融资。今年值得关注的新上榜公司有 4 家：Adept（专注领域：AI 模型开发）、Insitro（药物研发）、Runway（图像和视频编辑）、Synthesia（合成视频创作）。



五、国内高校动态

1. 教育部：全国 498 所高校成功申报人工智能专业²⁶

2023 年 4 月，教育部网站公布 2022 年度普通高等学校本科专业备案和审批结果，据专业备案结果显示，2022 年全国共有 58 所高校成功申报人工智能专业，专业代码 080717T，专业学习年限四年，授予工学学士学位。迄今为止，全国共有 498 所高校成功申报人工智能本科专业。第一批 2018 年，全国 35 所高校成功申报人工智能专业；第二批 2019 年，全国 180 所高校成功申报人工智能专业，第三批 2020 年全国 130 所高校成功申报人工智能本科专业，第四批 2021 年全国 95 所高校成功申报人工智能本科专业，为人工智能研究型、应用型人才培养奠定了基础。

²⁶ [全国 498 所高校成功申报人工智能专业 教育部公布名单\(2023 年\)| 网易订阅](#) ,4 月 18 日

2. 教育部：全国 71 所高校成功申报数字经济专业²⁷

2023 年 4 月，教育部网站公布 2022 年度普通高等学校本科专业备案和审批结果，据专业备案结果显示，2022 年全国共有 71 所高校成功申报数字经济专业，专业代码 020109T，专业学习年限四年，授予经济学学士学位。数字经济是以人工智能、大数据、区块链、云计算等底层数字技术驱动的。

3. 高校动态

清华大学



- 4 月 8 日，清华大学与蚂蚁集团签署合作协议，双方将在“下一代互联网应用安全技术”方向展开合作，聚焦智能风控、反欺诈等核心安全场景，携手攻坚可信 AI、安全大模型等关键技术，并加速技术落地应用，以解决 AI 时代的互联网安全科技难题，筑牢数字安全屏障。
- 4 月 8 日，清华大学社会科学学院积极心理学研究中心和全球产业研究院联合厦门大学社会与人类学院社会认知与计算实验室和合芯科技，共同举办“生成式人工智能与人类未来：机会与挑战”跨学科高峰论坛。本次论坛主题包括：生成式人工智能与产业革命、人类优势、教育变革、网络安全、文化偏差等。

北京大学



- 4 月 23 日，北京大学武汉人工智能研究院在武汉揭牌。北京大学武汉人工智能研究院还与浪潮新基建科技有限公司、中国联通智慧足迹数据科技有限公司等 8 家公司进行了战略伙伴签约，共同构建智能社会治理的“产-学-研-用”生态。

上海交通大学



- 4 月 2 日，上海市依托上海交通大学和上海白玉兰开源开放研究院布局建设的“AI for Science 科学数据开源开放平台”正式启动，该平台将助力上海占领 AI for Science（人工智能驱动的科学研究的）高地。
- 4 月 6 日，重庆智能工程职业学院与上海交通大学重庆人工智能研究院在永川举行合作签约仪式暨战略合作工作推进会。此次校院合作，双方将充分发挥各自在教育、科研及人才等方面的资源优势，在人才培养、

²⁷ [全国 71 所高校成功申报数字经济专业 教育部公布名单\(2023 年\)| 网易订阅](#)，4 月 19 日

科学研究、乡村振兴和科研平台建设等方面开展多层次、多形式的合作，

中国人民大学 ● 近日，快手与中国人民大学高瓴人工智能学院联合成立未来媒体智能联合实验室。双方将打通产学研链路，开展联合研究、多应用场景探索和人才培养合作，在多模态大模型、AIGC、智能搜索与推荐、数字人等多个人工智能前沿领域持续推进技术创新。



香港理工大学 ● 香港理工大学人工智能机器人实验室 4 日正式启用，以加强机器人科技教育，并促进学生及科研人员的共同创造和创新。该实验室已有 15 个理大项目团队进驻，配置先进机械臂、机器人、自动导引车、无人机，以及人工智能机器人和工业 4.0 相关的学习套件。



南开大学 ● 日前，由天津市教育委员会主办的“启诚杯”第 11 届天津市大学生学科竞赛人工智能电脑鼠大赛在南开大学开赛，共有来自天津大学、南开大学、天津工业大学等 16 所高校的 294 支参赛队 900 余名师生参加“古典电脑鼠”和“虚拟电脑鼠”两个赛道的比赛。



北京师范大学 ● 联合国教科文组织人工智能与教育教席近日在北京师范大学揭牌成立。教席将于今年上半年正式启动，并在 4 年任期内开展教育领域人工智能相关研究、培训、信息共享和文献工作，推动人工智能技术创新，探究有效且合乎伦理的人工智能教育应用，助力相关人才的培养。教席启动仪式后还举行了“人工智能与数字教育”研讨会，会上发布了《人工智能为所有人学习所用？100 篇有影响力的教育机器人学术论文》手册。



华中科技大学 ● 4 月 11 日，江夏区人民政府与华中科技大学签署联合建立“人工智能与自动化创新联合研究中心”合作协议。联合研究中心将围绕江夏区目前已有产业，依托华中科技大学现有技术，包括工业人工智能、微电网能量管理系统、智能制造等领域，联合开展关键技术攻关、行业标准制订和产业生态建设。



吉林大学 ● 4 月 24 日，智慧停车一体化平台捷停车与吉林大学数学学院达成战略合作，双方将联合发挥各自的人才资源优势 and 科研创新优势，共建“产学研基地”。深入开展包括“停车难”区域判定模型、停车场智能经营模式、充电桩运营优化策略等多项智慧停车相关的“人工智能+大数据”课题研究、新技术应用示范与人才培养等方面的合作。



复旦大学 ● 中国自动化学会、复旦大学、百度飞桨联合举办“人工智能与大模型”师资培训班，本次活动将重点阐释大模型的产品能力、技术原理和产业价值。对象包括各高校人工智能相关专业教师等。参训教师将获教育部



产学研合作协同育人项目优先支持。

南京大学



- 4月15日，南京大学计举行首届人工智能女科学家论坛，汇聚了国内人工智能相关的顶尖女性科学家，共同探讨人工智能领域的前沿技术与发展趋势。

浙江大学



- 4月28日，浙江大学教育基金会潘云鹤人工智能科教基金二期捐赠仪式举行。同时，第一届“潘云鹤人工智能科教奖”评选工作正式启动。

江南大学



- 江南大学近期成立了以世界著名人工智能专家 Josef Kittler 名字命名的 Josef Kittler 人工智能研究院，主要进行人工智能高水平人才培养、人工智能应用基础研究、人工智能应用技术开发和人工智能技术转移与转化。4月28日，无锡高新区与江南大学双方共建 Josef Kittler 人工智能研究院项目签约仪式。



六、AI 人才动态

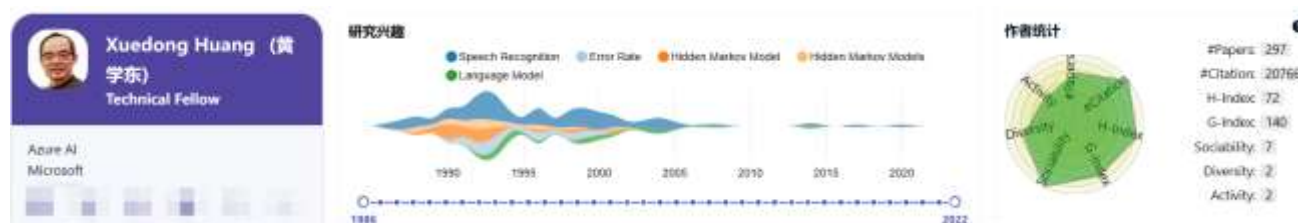
【学术界】

1. 微软黄学东入选美国艺术与科学院新增院士²⁸

近日，美国艺术与科学院（American Academy of Arts & Sciences，简称 AAAS）公布了新一批院士的名单，共有来自数学与物理科学、生命科学、社会与行为科学、人文艺术、公共事务与行政管理学领域的 269 人当选 2023 AAAS

²⁸ [微软黄学东、奥斯卡影后杨紫琼入选！2023 美国艺术与科学院院士名单公布](#)，4月20日

院士 (或外籍院士), 以表彰他们为领域发展做出的贡献。有多位华人学者当选, 其中, 计算机科学领域, 微软全球人工智能首席技术官黄学东入选, 他的研究方向为语音识别技术。



2. 陈怡然荣获 2022 Charles A. Desoer 技术成就奖²⁹

4 月 22 日, 杜克大学电子和计算机工程系的陈怡然教授荣获 2023 年 IEEE 电路与系统分会 Charles A. Desoer 技术成就奖。获奖者可以得到 2000 美元的奖金和证书。陈怡然的研究领域包括: 新型存储器及存储系统, 机器学习与神经形态计算, 以及移动计算系统等。2022 年, 该奖项由香港城市大学的谢智刚教授获得, 他的研究集中在复杂网络应用、电力电子、非线性系统和智能电力分配; 2021 年, 该奖项由南加州大学郭宗杰获得, 他的研究集中在数字媒体处理、多媒体压缩、通讯网络技术及嵌入式多媒体系统设计等领域。

²⁹ [3 年 3 位华人! 陈怡然荣获 2022 Charles A. Desoer 技术成就奖, 4 月 23 日](#)

【产业界】

3. Jeff Dean 将担任 Google DeepMind 的首席科学家³⁰

4月21日，Google CEO Sundar Pichai 宣布，将 DeepMind、Google Brain 合并，成立 Google DeepMind，以安全和负责任的方式建立能力更强的 AGI。Demis Hassabis 将担任 Google DeepMind 的首席执行官，Jeff Dean 将担任首席科学家一职。

4. 搜狗前 CEO 王小川官宣大模型创业，对标产品今年推出³¹

4月10日，搜狗创始人、前 CEO 王小川宣布开启 AI 大模型领域创业，与 COO 茹立云联合成立大语言模型公司百川智能。百川智能以前搜狗团队为基础，囊括来自百度等大厂的领军人物和创企核心人员，到这个月底将会有近 50 人规模的精英团队。公司已有 500 亿参数的大模型启动训练，年底前会推出一个对标 ChatGPT-3.5 的更大规模模型。



³⁰ [官宣，Google DeepMind 成立，4月21日](#)

³¹ [搜狗前 CEO 王小川官宣大模型创业！对标 GPT-3.5 的产品今年推出](#)，智东西，4月10日

5. IDEA 研究院张家兴入局大模型创业³²

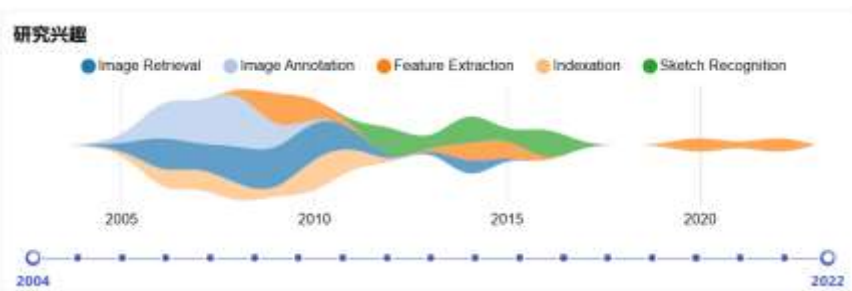
张家兴，粤港澳大湾区数字经济研究院（IDEA 研究院）讲席科学家、认知计算与自然语言中心负责人，被曝已经启动了创业项目，新公司据说代号深度空间，大方向是打造大模型底座，希望瞄准 B 端市场打造企业级应用。首轮估值已经高达 3 亿美元（折合人民币 20 亿）。知名 AI 大牛、IDEA 研究院创始理事长沈向洋担任顾问。张家兴，北京大学电子系博士毕业，师从侯士敏；本科阶段在南开大学学习嵌入式系统，主要研究领域为 AI、DL 和分布式系统；曾先后就职于百度、微软、阿里巴巴，曾任微软亚洲研究院研究员、360 数科首席科学家。张家兴参与过 IDEA 研究院发布的“悟道·文源”等大模型项目，带领团队开源了第一个以中文为主的 AI 绘画模型太乙。

6. 字节技术大牛创业打造聚焦 AIGC 的视觉多模态算法平台³³

字节跳动前视觉技术负责人、AI Lab 总监王长虎近日辞职进入 AIGC 创业，成立新公司爱诗科技，要打造一个聚焦 AIGC 的视觉多模态算法平台，覆盖视觉相关的各种垂直应用场景。目前组建了 10 余人的初始团队，在招募核心技术和产品人才。

³² [大模型新势力：IDEA 研究院张家兴入局创业，沈向洋任顾问，首轮估值 20 亿，4 月 25 日](#)

³³ [字节技术大牛跑步进入 AIGC 创业，聚焦视觉领域，搭建算法平台，还是多模态的，4 月 7 日](#)



7. 小米任命栾剑为 AI 实验室大模型团队负责人³⁴

近日小米集团发布内部邮件，任命栾剑担任技术委员会 AI 实验室大模型团队负责人，向技术委员会副主席、AI 实验室主任王斌汇报。栾剑现任小米技术委员会 AI 实验室语音生成团队负责人。曾任东芝（中国）研究院研究员、微软（中国）工程院高级语音科学家、微软小冰首席语音科学家及语音团队负责人等职位。目前小米 AI 团队超过 1200 人。



七、子领域重要科研进展

1. 机器学习

机构/个人名称	新闻动态	日期	链接
牛津大学	研究团队开发的框架 PureJaxRL 可以极大降低进入 Deep RL 研究的算力需求。FLAIR 研究人员分享了如何使用 JAX 框架仅利用 GPU 来高效运行强化学习算法，实现了超过 4000 倍的加速；并利用超高的性能，实现元进化发现算法，更好地理解强化学习算法。	4 月 14 日	详情

³⁴ 独家 | 小米任命栾剑为 AI 实验室大模型团队负责人 10%公司 澎湃新闻, 4 月 14 日

	<p>文章标题: <i>Achieving 4000x Speedups and Meta-Evolving Discoveries with PureJaxRL</i></p> <p>文章链接: https://chrislu.page/blog/meta-disco/</p>		
斯坦福大学	<p>Hazy Research 实验室的 Tri Dao 和 Dan Fu 等研究人员介绍了关于在高层级上增加序列长度的新方法, 并提供了连接一组新原语的“桥梁”。他们证明了 32k 的序列长度是可能的, 且在当前这个基础模型时代将得到广泛应用。</p> <p>论文标题: <i>FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness</i></p> <p>论文地址: https://arxiv.org/abs/2205.14135</p>	4 月 27 日	详情
香港科技大学	<p>开源项目 LMFlow 由港科大统计和机器学习实验室团队发起, 致力于建立一个全开放的大模型研究平台, 支持有限机器资源下的各类实验。基于 70 亿参数的 LLaMA, 只需 1 张 3090、耗时 5 个小时, 就可以训练一个专属于自己的个性化 GPT, 并完成网页端部署。LMFlow 拥有四大特性: 可扩展、轻量级、定制化和完全开源。</p> <p>lmflow.com 地址: http://lmflow.com/</p>	4 月 2 日	详情
	<p>港科大 LMFlow 团队提出全新对齐算法 RAFT (Reward rAnked Fine-Tuning)。RAFT 在训练过程中具有较高的稳定性, 能取得更好的对齐效果。并且任意生成模型都可以用此算法高效对齐, NLP/CV 通用。用在 Stable Diffusion 上, 还能对齐生成图片和提示词, 让模型生成更加符合提示词描述的图片。RAFT 的对齐训练过程中生成与训练过程完全解耦。</p>	4 月 17 日	详情
上海交通大学、牛津大学、亚马逊	<p>通过试图建立扩散微分方程与神经网络架构的联系, 研究人员阐释了能量约束扩散过程与各类信息传递网络的联系, 并为新的信息传递设计提供了一种理论参考。基于此, 提出了一种新型的可扩展 Transformer 模型, 称为 DIFFormer (diffusionbased Transformers)。它可以作为一种通用的 encoder, 在前向计算中利用样本间隐含的依赖关系。实验表明在小 / 大图节点分类、图片 /</p>	4 月 29 日	详情

	<p>文本分类、时空预测等多个领域的实验任务上 DIFFormer 都展现了强大的应用潜力；在计算效率上，只需要 3GB 显存就可以实现十万级样本间全联接的信息传递。</p> <p>论文标题: <i>Diffformer: Scalable (Graph) Transformers Induced By Energy Constrained Diffusion</i></p> <p>论文地址: https://arxiv.org/pdf/2301.09474.pdf</p>		
微软研究院	<p>微软研究院首次使用 GPT-4 作为教师模型进行 self-instruct 微调。研究者一方面发布了 GPT-4 生成的数据，包括中英文的 52k 指令遵循数据集、GPT-4 生成的对三种指令微调模型的输出进行评级的反馈数据。另一方面基于 GPT-4 生成的数据开发了指令微调的 LLaMA 模型和奖励模型。实验结果验证了使用 GPT-4 生成的数据进行 LLMs 指令微调的有效性。</p> <p>论文标题: <i>Instruction Tuning with GPT-4</i></p> <p>论文链接: https://arxiv.org/pdf/2304.03277.pdf</p>	4 月 9 日	详情
微软	<p>4 月 12 日，微软 (MSFT) 宣布开源 Deep Speed Chat，一个可以在模型训练中加入完整 RLHF 流程的系统框架，帮助用户训练类 ChatGPT 等大语言模型。Deep Speed Chat 基于微软 Deep Speed 深度学习优化库开发，具备训练、强化推理等功能，使用 RLHF(人工反馈机制的强化学习)技术，可将训练速度提升 15 倍以上，并大幅降低成本。</p> <p>项目地址: https://github.com/microsoft/DeepSpeed</p>	4 月 12 日	详情 1; 详情 2; 详情 3
微软亚洲研究院、韩国科学技术院 (KAIST)	<p>研究人员提出了首个适配所有密集预测任务的小样本学习器 VTM，以轻量化的迁移成本，赋予了计算机视觉模型预测新任务标签的能力，为计算机视觉中密集预测任务的处理以及小样本学习方法打开了全新思路。</p> <p>论文标题: <i>Universal Few-shot Learning of Dense Prediction Tasks with Visual Token Matching</i></p> <p>论文地址: https://arxiv.org/abs/2303.14969</p>	4 月 24 日	详情
OpenAI	<p>OpenAI 开源了模型 (Consistency Models)，在 GitHub 上公开了模型权重。该模型比扩散模型更快、性</p>	4 月 13 日	详情;

	能更好，无需对抗训练，就能生成高质量图片；测试显示，3.5 秒左右就能生成 64 张左右 256×256 的图像。 论文标题: <i>Consistency Models</i> 论文地址: https://arxiv.org/abs/2303.01469		详情 2
谷歌、哥伦比亚大学	谷歌助手团队和哥伦比亚大学俞舟教授合作推出 PRESTO 数据集，这是一个用于解析现实任务对话的多语言数据集，其中包括大约 50 万人与虚拟助理之间的现实对话。该数据集涵盖六种不同的语言，包括用户在使用助手时可能遇到的多种对话场景，包括用户定制改变、不流畅的对话语序场景、不同语言混合使用场景。数据集还包括结构化的上下文，例如用户的联系人列表。 PRESTO 对各种场景进行了标注，使人们能够创建不同的测试集来分别分析这些语言场景的模型性能。 论文名称: <i>PRESTO: A Multilingual Dataset for Parsing Realistic Task-Oriented Dialogs</i> 论文链接: https://arxiv.org/abs/2303.08954	4 月 7 日	详情
南方科技大学	李依明团队研发出一款名为 FD-DeepLoc 的深度学习框架，借此将单分子定位显微镜 (SMLM, Single-molecule localization microscopy) 的有效成像视野扩展到约 180×180μm ² 的范围，并在全细胞范围实现三维超高分辨成像，相比之前实现了高出 100 多倍的成像通量。该成果突破了传统卷积神经网络对于空间位置不敏感的局限，在大面阵科研级互补金属氧化物半导体相机的整个感光芯片视场范围内，均能对空间移变的单分子数据精准定位，并能实现三维大视场高通量超分辨成像。 论文标题: <i>Field-dependent deep learning enables high-throughput whole-cell 3D super-resolution imaging</i> 论文链接: https://www.nature.com/articles/s41592-023-01775-5	4 月 18 日	详情
北京智源人工智能研究院等	Chinese Open Instruction Generalist (COIG) 第一期共发布了 5 个子数据集，包括翻译指令、考试指令、人类价值观对齐指令、反事实修正多轮聊天、Leetcode 指令，总计 191k 数据，聚焦中文语料、数	4 月 21 日	详情

	<p>据类型多样、经过了人工质检与修正、数据质量可靠，而且可以商用。</p> <p>论文标题: <i>Chinese Open Instruction Generalist: a Preliminary Release</i></p> <p>论文链接: https://arxiv.org/pdf/2304.07987.pdf</p>		
DeepPavlov 等机构	<p>研究者通过采用一种 Recurrent Memory Transformer (RMT) 的架构，将 BERT 模型的有效上下文长度增加到 200 万个 token（约 3200 页文本），同时保持了较高的记忆检索准确性。新方法允许存储和处理局部和全局信息，并通过使用 recurrence 使信息在输入序列的各 segment 之间流动。</p> <p>论文标题: <i>Scaling Transformer to 1M tokens and beyond with RMT</i></p> <p>论文地址: https://arxiv.org/pdf/2304.11062.pdf</p>	4 月 25 日	详情
UC Austin、谷歌、香港中文大学等	<p>研究者提出了一种有效的对齐感知训练策 AligNeRF 可以提高 NeRF 的性能。该方法几乎没有引入明显的训练和测试成本，而在不同数据集上的实验表明，与基本的 NeRF 模型相比，该工作可以恢复更多的高频细节。</p> <p>论文标题: <i>AligNeRF: High-Fidelity Neural Radiance Fields via Alignment-Aware Training</i></p> <p>论文地址: https://arxiv.org/abs/2211.09682</p>	4 月 25 日	详情
纽约大学、谷歌、UBC、Vector Insstitute	<p>图灵奖得主 Geoffrey Hinton 等研究者提出了一系列新架构和算法改进，使得前向梯度学习对标准深度学习基准测试任务具有实用性。该研究表明，通过将扰动（perturbation）应用于激活而不是权重，可以显著减少前向梯度估计器的方差。研究团队通过引入大量局部贪心损失函数（每个损失函数只涉及少量可学习参数）和更适合局部学习的新架构 LocalMixer（受 MLP Mixer 启发），进一步提高了前向梯度的可扩展性。该研究提出的方法在 MNIST 和 CIFAR-10 上与反向传播性能相</p>	4 月 30 日	详情

	<p>当，并且明显优于之前 ImageNet 上的无反向传播算法。</p> <p>论文标题： <i>Scaling Forward Gradient With Local Losses</i></p> <p>论文链接： https://arxiv.org/abs/2210.03310</p>		
中国人民大 学	<p>20 余位研究者联合撰写大型语言模型综述，指出由于涌现能力与自然界中的相变有着密切的相似性，跨学科的理论或原则可能对解释和理解 LLMs 的行为有帮助；由堆叠的多头自注意力层组成的 Transformer 已经成为构建 LLMs 的普遍架构，但研究更高效的 Transformer 变体对构建 LLMs 的影响是很重要的，例如稀疏注意力已经被用于 GPT-3；到目前为止，仍然缺乏具有完整的预处理和训练日志的开源模型检查点以供重现的 LLM。</p> <p>论文名称： <i>A Survey of Large Language Models</i></p> <p>论文链接： https://arxiv.org/abs/2303.18223</p>	4 月 3 日	详情

2. 计算机视觉

机构/个人名称	新闻动态	日期	链接
Meta	<p>4 月 6 日，Meta AI 发布了首个图像分割基础模型 Segment Anything Model (SAM)。SAM 能从照片或视频中对任意对象实现一键分割，生成 mask，甚至包括在训练过程中没有遇到过的物体和图像类型。并且能够零样本迁移到其他任务。可以在一个统一框架 prompt encoder 内，指定一个点、一个边界框、一句话，直接一键分割出物体。</p> <p>论文标题： <i>Segment Anything</i></p> <p>论文链接： https://arxiv.org/abs/2304.02643</p> <p>论文地址： https://ai.facebook.com/research/publications/segment-anything/</p>	4 月 6 日	详情 ; 详情 2 ; 详情 3
Meta、腾讯	Meta AI 的基础人工智能研究 (FAIR) 团队发布了首	4 月 14	详情

美国、卡内基梅隆大学	<p>个 AI 动画绘图 (Animated Drawings) 开源项目, 用 AI 技术轻松让各种人物涂鸦变成动画。Meta AI 发布了该项目代码以及包含近 180000 张带标注的业余绘图的新数据集。</p> <p>论文名称: <i>A Method for Animating Children's Drawings of the Human Figure</i></p> <p>论文链接: https://arxiv.org/abs/2303.12741</p>	日	
Meta、UTAustin、Cruise	<p>研究者联合提出了开放语言风格模型 (open-vocabulary segmentation, OVSeg), 包括一个分割模型和一个 CLIP 模型, 它能让 Segment Anything 模型知道所要分隔的类别。</p> <p>论文名称: <i>Open-Vocabulary Semantic Segmentation with Mask-adapted CLIP</i></p> <p>论文链接: https://arxiv.org/pdf/2210.04150.pdf</p>	4月15日	详情
Meta、Inria	<p>研究人员收集了一个由 1.42 亿张图片组成的多样化的语料库, 最终提供了各种预训练的视觉基础模型 DINOv2。研究人员用 1B 的参数训练了一个 ViT 模型, 并将其提炼成一系列较小的模型, 这些模型在大多数图像和像素级别上超过了现有的 OpenCLIP 在图像和像素层面上的基准。DINOv2 模型能产生高性能的视觉表征, 无需微调就能用于分类、分割、图像检索、深度估计等下游任务。具有如下特征: (1) 使用自监督的方式进行训练, 而不需要大量的标记数据; (2) 可以用作几乎所有 CV 任务的骨干, 不需要微调; (3) 直接从图像中学习特征, 而不依赖文本描述, 这可以使模型更好地理解局部信息; (4) 可以从任何图像集合中学习; (5) DINOv2 的预训练版本已经可用, 并可以在一系列任务上媲美 CLIP 和 OpenCLIP。</p> <p>论文标题: <i>DINOv2: Learning Robust Visual Features without Supervision</i></p> <p>论文地址: https://arxiv.org/pdf/2304.07193.pdf</p>	4月18日	详情1; 详情2
谷歌	<p>谷歌研究人员提出了一种能够高效且稳定训练大规模 Vision Transformers (ViT) 模型的方法, 将 ViT 的参数量</p>	4月6日	详情

	<p>提升到 220 亿。ViT-22B 使用 QK 归一化改进了训练稳定性，提出了一种异步并行线性操作 (asynchronous parallel linear operations) 的新方法提升训练效率，并且能够在硬件效率更高的 Cloud TPU 上进行训练。ViT-22B 可以应用于 PaLM-e 中，与语言模型结合后的大模型可以显著提升机器人任务的技术水平。优势包括更好地平衡公平性和性能，在形状/纹理偏见方面与人类视觉感知的一致性，以及更好的稳健性。</p> <p>论文名称: <i>Scaling Vision Transformers to 22 Billion Parameters</i></p> <p>论文链接: https://arxiv.org/abs/2302.05442</p>		
	<p>研究员提出一个叫 Speed is all you need 的方法：通过 GPU 优化加速大规模扩散模型在设备上的推理速度，可以 12 秒内 AI 在手机上完成作。该方法是针对 Stable Diffusion 来做的优化，但同时也能适应其他扩散模型。面向的任务是从文本生成图像。</p> <p>论文标题: <i>Speed Is All You Need: On-Device Acceleration of Large Diffusion Models via GPU-Aware Optimizations</i></p> <p>论文地址: https://arxiv.org/abs/2304.11267</p>	4 月 27 日	详情
美国东北大学、麻省理工学院	<p>研究者提出了一种在预训练之后有选择地从文本条件模型的权重中删除单个概念的方法，不需要重新训练。</p> <p>论文名称: <i>Erasing Concepts from Diffusion Models</i></p> <p>论文链接: https://arxiv.org/pdf/2303.07345v1.pdf</p>	4 月 7 日	详情
新加坡国立大学	<p>学者们开发了一个 Anything-3D 框架来解决 2D 图像的 3D 重建问题。借助 Meta 的“分割一切”模型，Anything-3D 直接让分割后的任意物体活起来了。采用 BLIP 模型生成纹理描述，用 SAM 模型提取图像中的物体，然后利用文本→图像的扩散模型 Stable Diffusion 将物体放置到 Nerf（神经辐射场）中。另外，再用上 Zero-1-to-3 模型，可以得到不同角度的柯基，还可以进行人物 3D 重建。</p> <p>论文名称: <i>Anything-3D: Towards Single-view</i></p>	4 月 24 日	详情

	<p><i>Anything Reconstruction in the Wild</i></p> <p>论文链接: https://arxiv.org/pdf/2304.10261.pdf</p>		
威斯康辛麦迪逊、微软、港科大等	<p>研究人员提出 SEEM 模型, 通过不同的视觉提示和语言提示, 实现一次性分割所有地方的一切图像视频。在模型架构上, SEEM 采用了常见的编码器-解码器架构。其独特的地方在于具有查询和提示之间复杂的交互。</p> <p>论文名称: <i>Segment Everything Everywhere All at Once</i></p> <p>论文链接: https://arxiv.org/pdf/2304.06718.pdf</p>	4月23日	详情
英伟达	<p>英伟达提出一种可对未知物体的 6D 姿态追踪和三维重建的方法 BundleSDF。用于从单目 RGBD 视频序列中跟踪未知物体的 6 自由度运动, 同时进行物体的隐式神经三维重建, 方法接近于实时 (10Hz)。这种方法适用于任意刚性物体, 即使视觉纹理大部分确实, 仅需在第一帧中分割出物体, 不需要任何额外的信息, 并且不对智能体与物体的交互模式做任何假设。</p> <p>论文标题: <i>BundleSDF: Neural 6-DoF Tracking and 3D Reconstruction of Unknown Objects</i></p> <p>项目地址: https://bundlesdf.github.io/</p>	4月2日	详情
英伟达、慕尼黑大学等	<p>近期学者联合发布了一篇有关视频潜伏扩散模型 (Video LDM) 的论文, 该模型可将文本转换成视频。研究者利用潜在扩散模型 (latent diffusion model, LDM) 实现了高分辨率的长视频合成, 做到最高 1280×2048、最长 4.7 秒。研究者提出了视频潜在扩散模型 (Video LDM), 并将 LDM 扩展到了计算密集型任务: 高分辨率视频生成。在 LDM 的基础上, 本方法以计算和内存高效的方式生成了全局连贯的长视频。</p> <p>论文标题: <i>Align your Latents: High-Resolution Video Synthesis with Latent Diffusion Models</i></p> <p>论文地址: https://arxiv.org/abs/2304.08818</p>	4月20日	详情
智源研究院	<p>视觉团队推出通用分割模型 SegGPT (Segment Everything In Context), 利用视觉提示 (prompt) 完</p>	4月8日	详情

	<p>成任意分割任务的通用视觉模型。SegGPT 是智源通用视觉模型 Painter (CVPR 2023) 的衍生模型，针对分割一切物体的目标做出优化。SegGPT 训练完成后无需微调，只需提供示例即可自动推理并完成对应分割任务，包括图像和视频中的实例、类别、零部件、轮廓、文本、人脸等等。具有以下优势能力：通用能力、通用能力、自动视频分割和追踪能力</p> <p>论文名称： <i>SegGPT: Segmenting Everything In Context</i></p> <p>论文链接： https://arxiv.org/abs/2304.03284</p>		
美图影像研究院（MT Lab）、中国科学院大学	<p>研究提出一种新颖且即插即用的正则化器 DropKey，可以有效缓解 Vision Transformer 中的过拟合问题。该方法通过在注意力计算阶段随机 drop 部分 Key 以鼓励网络捕获目标对象的全局信息，从而避免了由过于聚焦局部信息所引发的模型偏置问题，继而提升了基于 Transformer 的视觉类算法的精度。</p> <p>论文名称： <i>DropKey for Vision Transformer</i></p> <p>论文链接： https://arxiv.org/abs/2208.02646</p>	4月19日	详情
天津大学、英国卡迪夫大学	<p>联合团队提出一种兼顾精细语义与几何刻画的三维人体表示 SemanticHuman，可实现部件级别的灵活可控编辑。作者通过提出一种部件感知的骨骼分离的解耦策略，在隐编码和人体部件的几何测量之间建立起明确的对应关系，不仅为表示提供了细粒度的语义、高效且强大的几何刻画能力，而且方便设计鲁棒的无监督损失，从而摆脱对配对监督数据的依赖。借助该表示所学习到的具有几何意义的隐空间，用户可以通过调节隐编码实现人体生成与编辑、人体形状风格转移等一系列有趣的应用，相关代码已经开源。</p> <p>论文名称： <i>Learning Semantic-Aware Disentangled Representation for Flexible 3D Human Body Editing</i></p> <p>论文链接： http://cic.tju.edu.cn/faculty/likun/projects/SemanticHuman</p>	4月18日	详情
西安交通大学等	<p>研究人员提出了 SadTalker (Stylized Audio-Driven Talking-head) 模型，在三维运动场中进行学习从音频中生成 3DMM 的 3D 运动系数（头部姿势、表情），并使用</p>	4月19日	详情

	<p>一个全新的 3D 面部渲染器来生成头部运动。音频可以是英文、中文、歌曲，视频里的人物还可以控制眨眼频率。</p> <p>论文名称: <i>SadTalker: Learning Realistic 3D Motion Coefficients for Stylized Audio-Driven Single Image Talking Face Animation</i></p> <p>论文链接: https://arxiv.org/pdf/2211.12194.pdf</p>		
Midjourney	<p>近日 推出了 Niji-journey V5, 它是一个针对二次元动漫的版本。Niji V5 支持了风格化 (stylize) 参数设置, 成图效果有时更写实、看起来更像真人</p>	4月13日	详情
阿里巴巴	<p>研究者专注于新的视图合成任务, 提出了一个名为 4K-NeRF 的框架, 其基于 NeRF 的体积渲染方法可以实现在 4K 超高分辨率下高保真视图合成, 增强其在以极高分辨率的场景中恢复视图一致的细微细节的表现力。该研究还引入了一对保持几何一致性的编解码器模块。该研究希望将框架合并到动态场景建模中的效果, 以及神经渲染任务作为未来的方向。</p> <p>论文名称: <i>4K-NeRF: High Fidelity Neural Radiance Fields at Ultra High Resolutions</i></p> <p>论文链接: https://arxiv.org/abs/2212.04701</p>	4月14日	详情
OPPO	<p>OPPO 等机构的研究者们提出了一个通用学习框架 GlyphDraw, 旨在让模型能够生成嵌入连贯文本的图像, 这是图像合成领域首个解决汉字生成问题的的工作。该研究提出了首个汉字图像生成框架 GlyphDraw; 提出了一种有效的训练策略, 限制了预训练模型中可训练参数的数量, 以防止过拟合和灾难性遗忘; 提出了一个新的基准来使用 OCR 模型评估汉字图像生成质量。</p> <p>论文名称: <i>GlyphDraw: Learning to Draw Chinese Characters in Image Synthesis Models Coherently</i></p> <p>论文链接: https://arxiv.org/abs/2303.17870</p>	4月10日	详情

3. 自然语言处理

机构/个人名称	新闻动态	日期	链接
Meta	Meta 正在开发一个系统来读取人脑中的想法，该项目利用 AI 来帮助科学家们研究大脑中神经元，通过一个仿真人脑模型，实现人的大脑活动在无需打字、语音输入的情况下直接转文字。研究人员利用脑电图和脑磁图两种非侵入式技术获取数据，让自主监督学习 AI 工具 wave2vec 2.0 在开源录音上进行训练，训练数据集中包含 169 名志愿者在听到有声读物和单句时的大脑活动录音。通过近 150 个小时的录音，算法像懂读心术一样能够推断出人们最有可能听到的词。	4 月 3 日	详情
斯坦福大学、谷歌	<p>研究团队创造了 25 个 AI NPC，每个 NPC 都有不同的身份和行动决策，并让它们在一个沙盒环境“虚拟小镇”中共同生活，背后依靠 ChatGPT 大模型来完成行动决策。结果发现，这些 AI NPC 不仅会像人一样生活，而且彼此之间还会发生交互，甚至产生对话。要是分两步：一是构建各种各样的 AI “人设”，并让他们做出符合设定的日常行动；二是要让 AI 人设对外界环境和其他人的行动做出反应，产生交互。随后，无论是 AI 计划要做的事（plan）、还是对外界发生的事情给出反馈（reflect），都会受到这个记忆流的影响，具体方式是通过从记忆中检索对应的事件，并依靠大模型 ChatGPT 来进行决策。研究者们还介绍了 Generative Agents（生成式智能体），一种利用生成模型来模拟可信人类行为的智能体，并证明它们能产生可信的个人和突发群体行为的模拟。</p> <p>论文名称： <i>Generative Agents: Interactive Simulacra of Human Behavior</i></p> <p>论文链接： https://arxiv.org/abs/2304.03442 https://arxiv.org/pdf/2304.03442v1.pdf</p>	4 月 11 日	详情 1 ; 详情 2
摩智能、腾讯 AI Lab、华中	研究人员发现量化大型语言模型的主要挑战来自于通道之间不同的激活范围，而不仅仅是离群值问题。作	4 月 9 日	详情

科技大学、北京大学、伊利诺伊理工学院	<p>者提出了一种基于重排序的量化方法 RPTQ，解决了量化大型语言模型激活的问题。RPTQ 通过重新排列激活中的通道，然后将它们分簇进行量化，从而降低通道范围差异的影响。此外，作者通过避免显式重新排序来减少存储和计算开销。该工作首次将 LLM 模型的激活量化推向 3 比特。激活值的 3 比特量化同时也能与权重的 3 比特量化共同使用，大幅度降低了大型语言模型的开销。</p> <p>论文名称: <i>RPTQ: Reorder-based Post-training Quantization for Large Language Models</i></p> <p>论文链接: https://arxiv.org/abs/2304.01089</p>		
UC 伯克利、CMU、斯坦福	<p>团队正式发布了 130 亿参数的 Vicuna (小羊驼) 的权重，只需单个 GPU 就能跑。Vicuna 是通过在 ShareGPT 收集的用户共享对话上对 LLaMA 进行微调训练而来，训练成本近 300 美元。就能实现 ChatGPT 90%的性能</p>	4 月 4 日	详情
微软	<p>研究者从人类解决数学问题的方式中获得灵感，将其分解为更简单的多步骤程序，并在每个步骤中利用多种方式来验证他们的方法。使用了 175B 参数量的大型语言模型 GPT3 DaVinci completion engine，能够将模型在 MultiArith 数据集上的准确率从 78.7% 提升到 92.5%。</p> <p>论文名称: <i>MathPrompter: Mathematical Reasoning Using Large Language Models</i></p> <p>论文链接: https://arxiv.org/pdf/2303.05398.pdf</p>	4 月 7 日	详情
斯坦福大学	<p>开发者发布了一个 LLM 引擎 Lamini，可以为开发人员提供从 GPT-3 带到 ChatGPT 的超能力，使用很多公司、机构的基础模型快速构建定制化模型。Lamini 提供了一种托管化的数据生成器，只需执行 Lamini 库中的几行代码，用户就能训练自己的大型语言模型 (LLM) 及其权重，而无需使用任何 GPU。</p> <p>链接: https://lamini.ai/</p>	4 月 29 日	详情
罗格斯大学	<p>研究团队开源了一个 AGI 研究平台 OpenAGI，专</p>	4 月 26 日	详情

	<p>门用于提供复杂的、多步骤的任务，并辅助有特定任务的数据集、评估指标和各种可扩展的模型。此外，还提出了一个从任务反馈中强化学习 (RLTF) 的机制。研究人员认为，语言模型操作各种专家模型解决复杂任务的范式是实现 AGI 的一个很有前途的方法。</p> <p>论文标题: <i>OpenAGI: When LLM Meets Domain Experts</i></p> <p>论文链接: https://arxiv.org/pdf/2304.04370.pdf</p>	日	
北京大学	<p>团队提出了 Learning to Program (LP)方法，让大语言模型 LLM 从自然语言程序数据集中进行学习，并用学到的内容指导其推理过程。</p> <p>论文标题: <i>Learning to Program with Natural Language</i></p> <p>论文链接: https://arxiv.org/abs/2304.10464</p>	4 月 25 日	详情
OpenAI	<p>团队负责人 Jack Rae 提出：大语言模型的本质是一个性能强大的数据无损压缩器。语言模型的本质，是不断预测下一个词的概率分布，从而完成生成式任务。随着模型规模的提升，基于信息熵计算出的压缩率也会逐渐提升。更好的算法架构、基于 Plugin 的工具集成、合成数据的运用都能够有效提升模型的压缩率，从而实现模型性能的进一步提升。</p>	4 月 9 日	详情
Karpathy	<p>特斯拉前 AI 总监，刚刚回归 OpenAI 的 Andrej Karpathy 推出一种最简 GPT 的玩法。它是一个带有两个 token 0/1 和上下文长度为 3 的极简 GPT，将其视为有限状态马尔可夫链。通过简化，Karpathy 已让 GPT 模型变得易于可视化。</p>	4 月 10 日	详情
斯坦福大学、Ontocord.AI、苏黎世联邦理工学院等	<p>团队宣布开启“红睡衣” (RedPajama) 计划，旨在生成可复现、完全开放、最先进的语言模型，即从零一直开源到 ChatGPT。计划共包括三部分：高质量、大规模、高覆盖度的预训练数据集；在预训练数据集上训练出的基础模型；指令调优数据集和模型，比基本模型更安全、可靠。</p>	4 月 22 日	详情
宾夕法尼亚州	<p>研究人员发布关于指令学习的全面综述，探讨了自</p>	4 月 3 日	详情

立大学、俄亥俄州立大学	然语言处理中新兴的指令学习范式，重点关注任务指令、建模方法、性能影响因素及挑战。 论文名称: <i>Is Prompt All You Need? No. A Comprehensive and Broader View of Instruction Learning</i> 论文链接: https://arxiv.org/pdf/2303.10475v2.pdf	日	
-------------	--	---	--

4. AIGC/生成式 AI

(1) 基础模型

企业机构

机构/个人名称	新闻动态	日期	链接
阿里达摩院	达摩院在 ModelScope 上开源了“基于文本的视频生成”的模型参数。只需要输入文本，即可返回符合文本描述的视频，并且该模型能够适用于“开放领域”的视频生成，能够基于文本描述进行推理，目前只支持英文输入。	4月2日	详情
Hugging Face Inc.	开源社区 Hugging Face 发布了 HuggingChat ，有 300 亿参数，既开源又免费。 Open Assistant 和 HuggingChat 是两个相关的项目，都涉及使用深度学习和海量文本数据开发对话式人工智能。它们的共同根源是 Hugging Face Inc.。此外，HuggingChat 代表的是聊天机器人或语言模型等方面，而 Open Assistant 则建立在这一基础之上，具有额外的特性和功能，旨在更好地适用于更一般化的问题和答案。 网址: hf.co/chat	4月26日	详情
智谱 AI	智谱 AI GLM-130B 系列大模型目前与 GPT-3 基座模型对标； ChatGLM-6B 只用单个 GPU 就能跑起来的，全球下载量超过 100 万；辅助编程工具 CodeGeeX	4月21日	详情 1 ； 详情 2

	<p>每天帮助程序员“编写”超过 400 万行代码。</p> <p>在产品应用方面，智谱 AI 在自研的千亿参数大模型 GLM-130B 基础上，为上上签打造了签约智能产品 Hubble 哈勃。这是电子签约行业首款应用了大模型技术的生成式 AI 产品，可在用户授权下，通过连续对话实现对电子签约合同的检索，执行概况解读、重点识别、筛选标识、分类归纳等相关功能，帮助用户降本增效。</p>		
Stability AI	<p>官宣正在研发的 Stable Diffusion XL 模型开始面向公众测试，目前可以在 Clipdrop 平台免费试用。据称 SD-XL 是一个与 v2 模型架构相似、但规模和参数量更大的模型，大约有 23 亿参数。SD-XL 可以更好地理解输入的 prompt，并生成更精确的图像。</p> <p>试用链接：https://clipdrop.co/stable-diffusion</p>	4 月 17 日	详情
	<p>Stability AI 公司正式推出类 ChatGPT 产品 StableLM。目前处于 Alpha 版本，拥有的参数量分别为 30 亿和 70 亿；后续还将推出 150 亿到 650 亿参数模型。该模型基于 GPT-J、GPT-NeoX 和 Pythia 套件开发，在 The Pile 开源数据集上进行训练。能对话、创作内容、也能生成代码。</p> <p>GitHub 地址：https://github.com/Stability-AI/StableLM</p>	4 月 20 日	详情 1 ； 详情 2
	<p>StabilityAI 旗下的独立研发团队 DeepFloyd AI Research 推出了全新开源模型 DeepFloyd IF，目前以非商用许可开源。DeepFloyd IF 不仅图像生成质量高，还能准确绘制文字以及准确理解空间关系。可以把指定文字巧妙放置在画面中任何地方。DeepFloyd IF 基于扩散模型，负责理解文字的部分从 OpenAI 的 CLIP 换成了谷歌 T5-XXL，结合超分辨率模块中额外的注意力层，获得更准确的文本理解。负责生成图像的部分从潜扩散模型换成了像素级扩散模型。</p> <p>链接地址：https://huggingface.co/spaces/DeepFloyd/IF</p>	4 月 29 日	详情
	<p>Stability AI 在语言模型方面推出了首个开源并引入 RLHF 技术的聊天机器人 StableVicuna，基于小羊驼 Vicuna-13B 模型实现。目前代码和模型权重已开放下载。</p>	4 月 29 日	详情

	链接地址: https://huggingface.co/spaces/CarperAI/StableVicuna		
阿里云	阿里云大模型“ 通义千问 ”宣布开启企业邀测, 官网地址: tongyi.aliyun.com , 目前用户可以申请候补或者使用邀请码体验。具备包括文字创作、代码撰写、翻译等多语言能力。	4月7日	详情 ; 详情 2
商汤科技	商汤正式发布自研类 ChatGPT 产品: 商量 (SenseChat)。商量 依托的是商汤自研的千亿参数自然语言处理模型。此模型的背后, 则是商汤最新推出的日日新 (SenseNova) 大模型体系。商汤还发布了: 秒画 SenseMirage : AI 内容创作社区平台; 如影 SenseAvatar : AI 数字人视频生成平台; 琼宇 SenseSpace 、 格物 SenseThings : 3D 内容生成平台。这些 AIGC 产品所要发力之处, 将是医疗、短视频、教育、营销以及开发等产业领域。	4 月 10 日	详情 1 ; 详情 2
360	4 月 9 日, 360 官方正式宣布, 基于 360GPT 大模型开发的人工智能产品矩阵“360 智脑”率先落地搜索场景, 将面向企业用户开放内测。未来, 360 智脑将与浏览器、数字助理、智能营销等场景应用结合。从演示效果来看, 360 版 GPT 已经具备了一定的推理能力, 但在回答具体问题多次“翻车”。	4 月 10 日	详情
知乎、面壁智能	通过联合研发与战略投资的方式, 4 月 13 日知乎正式发布“ 知海图 AI ”中文大模型。背后的大语言模型 CPM-Bee 来自面壁智能。目前知海图 AI 模型已经运用进了知乎热榜, 知乎首个大模型功能“热榜摘要”上线并开启内测。未来, “知海图 AI”的能力将逐步应用在对创作的赋能、对讨论场的赋能和对信息获取的赋能。	4 月 13 日	详情 1 ; 详情 2
昆仑万维、奇点智源	4 月 17 日, 昆仑万维正式发布千亿级大语言模型“ 天工 ”, 同时宣布即日起启动邀请测试。该模型由昆仑万维与奇点智源联合研发。“天工”可通过自然语言与用户进行问答式交互, 其 AI 生成能力可满足文案创作、知识问答、逻辑推演、数理推算、代码编程等多元化需求。从“天工”目前发布版本来看, 最高已能支持 1 万字以上文本对话, 实现 20 轮次以上用户交互。	4 月 17 日	详情

	官网链接: tiangong.kunlun.com		
出门问问	<p>4 月 20 日, 出门问问的自研大模型 “序列猴子” (Sequence Monkey) 正式登场。基于这个大模型, 出门问问也推出了面向创作者的一站式 CoPilot 的 AIGC 产品矩阵, 具体包含四款产品: AI 写作平台 “奇妙文”、AI 绘画平台 “言之画”、AI 配音平台 “魔音工坊”、数字人视频与直播平台 “奇妙元”。目前四款产品开放注册, 全部直接可用。</p> <p>体验地址: write.mobvoi.com</p>	4 月 24 日	详情
百度	文心一言上线内测后, 一个月共迭代 4 次; 模型推理效率提升 10 倍; 推理性能提升 50%, 模型算力利用率提升 1 倍。截止目前, 已迭代至 3.5 版本。	4 月 24 日	详情
Anthropic	ChatGPT 最强对手产品 Claude 已经上线 Slack 平台, 采用了 Anthropic 称之为 “宪法式 AI” 的训练技术。每个人都可以免费使用。国内用户可以正常使用。	3 月 31 日	详情
Cerebras	这家 AI 芯片公司开源了 7 个 GPT 模型, 参数量分别达到 1.11 亿、2.56 亿、5.9 亿、13 亿、27 亿、67 亿和 130 亿。开放的模型不仅包含数据集, 可用于研究也可商用, 而且关键是预训练模型权重开放。	4 月 2 日	详情
超对称技术公司、复旦大学	超对称公司和复旦大学知识工场实验推出了 BBT-2, 全称 BigBang Transformer-2 (乾元 2), 120 亿参数的通用大模型。BBT-2 是基于 GPT Decoder-only 架构的大模型。项目团队还训练出了代码、金融、文生图等专业模型。	4 月 22 日	详情
科大讯飞	科大讯飞版 ChatGPT 产品 讯飞星火认知大模型 (Spark Desk) 对外开启内测。	4 月 25 日	详情
Databricks	4 月 12 日, Databricks 发布了 Dolly 2.0。Dolly 2.0 是业内第一个开源、遵循指令的 LLM, 它在透明且免费提供的数据集上进行了微调, 该数据集也是开源的, 可用于商业目的。基于 Dolly 2.0 模型, 用户可以修改和改进训练数据。	4 月 12 日	详情
白海科技	中文多模态模型 IDPChat, 可以生成图像文字, 只需 5 步+单 GPU。该模型以 IDP 平台为工具支撑, 以预训练	4 月 16 日	详情

	大语言模型 LLaMA 和开源文生图预训练模型 Stable Diffusion 为基础构建。目前可以同时支持文字对话和图片生成。 项目地址: https://github.com/BaihaiAI/IDPChat		
--	--	--	--

学术科研机构/官方政府

机构/个人名称	新闻动态	日期	链接
UC 伯克利	伯克利人工智能研究院 (BAIR) 发布了一个可以在消费级 GPU 上运行的对话模型 Koala (考拉)。Koala 使用从网络收集的对话数据对 LLaMA 模型进行微调。 Koala 模型在 EasyLM 中使用 JAX/Flax 实现, 并在配备 8 个 A100 GPU 的单个 Nvidia DGX 服务器上训练 Koala 模型。完成 2 个 epoch 的训练需要 6 个小时。在公共云计算平台上, 进行此类训练的成本通常低于 100 美元。Koala 已经推出线上测试 demo。Koala 还只是一个研究原型, 在内容、安全性和可靠性方面仍然存在重大缺陷, 也不应用于研究之外的任何用途。 项目地址: https://bair.berkeley.edu/blog/2023/04/03/koala/	4月5日	详情 ; 详情 2
浙江大学、北京大学、卡内基梅隆大学、中国人民大学	研究人员提出了音频理解与生成系统 AudioGPT 。AudioGPT 以 ChatGPT 充当负责对话与控制的大脑进行任务分析, 语音基础模型协同以完成跨模态转换、以及音频 (语音、音乐、背景音、3D 说话人) 模态的理解、生成, 能够解决 20 + 种多语种、多模态的 AI 音频任务。 论文标题: <i>AudioGPT: Understanding and Generating Speech, Music, Sound, and Talking Head</i> 论文地址: https://arxiv.org/pdf/2304.12995.pdf	4月30日	详情
微软亚研究院、浙江大学	研究团队发布了一个大模型协作系统 HuggingGPT 。该系统利用 ChatGPT 作为控制器, 连接 HuggingFace 社区中的各种 AI 模型, 来完成多模态复杂任务。到目前为止, HuggingGPT 已经涵盖了文本分类、	4月2日	详情

	<p>目标检测、语义分割、图像生成、问答、文本到语音、文本到视频等 24 个任务。</p> <p>论文名称: <i>HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face</i></p> <p>论文链接: https://arxiv.org/pdf/2303.17580.pdf</p>		
微软亚洲研究院	<p>近期推出多模态生成模型 NUWA-XL (eXtremely Long), 采用 Diffusion over Diffusion 架构, 通过“从粗到细”的生成过程, 以相同的粒度并行生成视频, 并应用全局扩散模型来生成整个时间范围内的关键帧, 然后通过局部扩散模型递归地填充附近帧之间的内容, 既提升了生成效率, 也确保了视频的质量和连续性。</p> <p>论文标题: <i>NUWA-XL: Diffusion over Diffusion for eXtremely Long Video Generation</i></p> <p>论文地址: https://arxiv.org/abs/2303.12346</p>	4 月 19 日	详情
复旦大学	<p>新版 MOSS 模型 21 日正式上线开源。相关代码、数据、模型参数已在 Github 和 Hugging Face 等平台开放。该项目所含代码采用 Apache 2.0 协议, 数据采用 CC BY-NC 4.0 协议, 模型权重采用 GNU AGPL 3.0 协议。MOSS 是一个支持中英双语和多种插件的开源对话语言模型, moss-moon 系列模型具有 160 亿参数, 在 FP16 精度下可在单张 A100 / A800 或两张 3090 显卡运行, 在 INT4/8 精度下可在单张 3090 显卡运行。MOSS 基座语言模型在约七千亿中英文以及代码单词上预训练得到, 后续经过对话指令微调、插件增强学习和人类偏好训练具备多轮对话能力及使用多种插件的能力。</p> <p>公开平台: https://moss.fastnlp.top/</p>	4 月 21 日	详情
全球非盈利机器学习研究机构 LAION	<p>LAION 发布了一个开源的对话模型 OpenAssistant。OpenAssistant 是基于 Pythia 和 LLaMA 微调而来, 主要用于训练人类标注的数据。这是一个由人工生成、人工标注的对话语料库, 包含 161,443 条消息, 分布在 66,497 个对话树中, 使用 35 种不同的语言, 并标注了 461,292 个质量评分。OpenAssistant Conversations 数据集是通过 13000 多名志愿者的众包努</p>	4 月 16 日	详情 1 ; 详情 2

	力综合而来的，英语和西班牙语占比最多，中文占 2.5%。		
俄罗斯联邦 储蓄银行	俄罗斯联邦储蓄银行公布 GigaChat ，对标 ChatGPT。GigaChat 最初将处于测试模式，仅限受邀请用户参与。与目前其他人工智能相比，GigaChat 最大的优势在于，能够用俄语进行比其他外国神经网络更智能的交流。	4 月 24 日	详情
阿卜杜拉国王科技大学	研究团队开源了 GPT-4 的平民版 MiniGPT-4 ，从效果上看已经基本具备 GPT-4 所展现出的多个功能，包括手写网页示意图生成代码、解读图像中的梗等。提出了一个两阶段的训练方法，先在大量对齐的图像-文本对上对模型进行预训练以获得视觉语言知识，然后用一个较小但高质量的图像-文本数据集和一个设计好的对话模板对预训练的模型进行微调，以提高模型生成的可靠性和可用性。 论文标题： <i>MiniGPT-4: Enhancing Vision-language Understanding with Advanced Large Language Models</i> 论文链接： https://arxiv.org/abs/2304.10464	4 月 21 日	详情

(2) 垂直领域基础模型

机构/个人名称	新闻动态	日期	链接
哈佛大学、斯坦福大学、耶鲁医学院、多伦多大学等	研究人员在 Nature 上联合发布提出了一种全新的医学人工智能范式，即“全科医学人工智能”（generalist medical artificial intelligence, GMAI）。 GMAI 模型 将能够使用很少或不使用任务特定的标记数据执行各种各样的任务。通过对大型、多样化数据集的自监督训练，GMAI 可以灵活地解释医学模式的不同组合，包括来自成像、电子健康记录、实验室结果、基因组学、图表或医学文本的数据。模型还可以生成具有表达能力的输出，如自	4 月 17 日	详情

	<p>由文本解释、口头推荐或图像注释，展现先进的医学推理能力。</p> <p>论文名称: <i>Foundation models for generalist medical artificial intelligence</i></p> <p>论文链接: https://www.nature.com/articles/s41586-023-05881-4</p>		
彭博社、约翰霍普金斯大学	<p>彭博社发布金融界的大型语言模型 BloombergGPT，拥有 500 亿个参数，专门针对金融领域的自然语言处理（NLP）任务进行训练，可以快速分析金融数据，帮助进行风险评估、情感分析、问答等功能。该模型是基于开源的 GPT-3 框架进行扩展和改进的，利用了彭博社丰富的金融数据源，构建了一个 3630 亿个标签的数据集，支持金融行业内的各类任务。该模型在多个标准的 LLM 基准测试和开放的金融基准测试上都超过了其他模型，在一系列内部开发的金融基准测试上也表现出了较高的水平。</p> <p>论文标题: <i>BloombergGPT: A Large Language Model for Finance</i></p> <p>论文链接: https://arxiv.org/abs/2303.17564</p>	3 月 31 日	详情 1 ; 详情 2
FATE 开源社区	<p>4 月 13 日，联邦学习隐私计算开源平台 FATE (Federated AI Technology Enabler) 正式上线发布新版本 FATE v1.11，推出了联邦大模型 FATE-LLM 功能模块。基于此技术方案，多个企业可以通过 FATE 内置的预训练模型如 GPT-2 进行横向联邦，利用各自隐私数据进行联邦大模型微调。过程中使用了安全聚合机制对各家模型数据进行保护。相对单一企业有限训练样本，通过联邦大模型技术综合多家的训练样本，可以显著提升模型效果和稳健性。经过实践测试，FATE-LLM FATEv1.11 集成了业界主流的大模型，为用户提供多种选择，其中包含 GPT-2、BERT、RoBERTa 等常用大语言模型，可以支持至少 30 家参与方同时进行横向联邦。</p> <p>项目地址: https://github.com/FederatedAI/FATE/releases/tag/v1.11.0</p>	4 月 13 日	详情 ; 详情
清华大学	<p>智能产业研究院（AIR）开源了一个生物医药研发领域的轻量级科研版基础模型 BioMedGPT-1.6B。参数 16 亿，最大特点是跨模态与知识融合。训练数据</p>	4 月 21 日	详情

	<p>中，包含分子、文献、专利、知识库等多尺度跨模态的生物医药大数据，并融合分子结构、知识图谱和文献文本中的知识，用于增强模型的泛化能力和可解释性。应用任务上，BioMedGPT-1.6B 可以处理药物性质预测、自然语言类、跨模态等多种任务。项目负责人聂再清。此次开源的 BioMedGPT-1.6B，其实是 BioMedGPT 的单机轻量版，后者是一个适用于生物医药领域研发的通用大模型。</p> <p>开源地址：https://github.com/BioFM/OpenBioMed</p>		
剑桥大学	<p>剑桥大学语言技术实验室官宣并开源一个带视觉的多模态医学基础大模型 Visual Med-Alpaca。这个专为生物医学领域设计的基础模型，构建于 Meta 开源的 LLaMa-7B 之上，还在一定程度上借鉴了 Visual-ChatGPT 和 Alpaca（羊驼）。因为参数量相对较小，所以使用单个消费级 GPU 就能轻松部署和运行。工作原理，是使用 prompt 管理器，将文本和视觉两种模态的信息合并到 prompt 中，从而生成具有生物医药专业知识的回答。该模型声明，严格用于学术研究。</p> <p>GitHub：https://github.com/cambridgeltl/visual-med-alpaca/tree/main/code</p>	4 月 15 日	详情
哈尔滨工业大学	<p>研究团队通过医学知识图谱和 GPT3.5API 构建了中文医学指令数据集，对 LLaMa 模型进行了指令微调构建并开源了医学智能问诊大模型 华佗: (HuaTuo)。相比于未经过医学数据指令微调的原 LLaMa 而言，HuaTuo 模型在智能问诊层面表现出色，可生成一些更为可靠的医学知识回答。与此同时，基于相同医学数据，该项目还训练了医疗版本的 ChatGLM 模型: ChatGLM-6B-Med。该团队还将发布扁鹊模型 PienChueh(同为基于医学数据训练的大模型)。</p> <p>论文标题: <i>HuaTuo: Tuning LLaMA Model with Chinese Medical Knowledge</i></p> <p>论文链接: https://arxiv.org/pdf/2304.06975v1.pdf</p>	4 月 29 日	详情
沃丰科技	<p>沃丰科技推出原心引擎 2.0，它从自有营销服务大数据资产出发，主打基于行业个性化和企业定制化的垂类大</p>	4 月 22 日	详情

	语言模型。原心引擎 2.0 基于 GPT 通用大模型，增量训练了多个行业的垂类数据，产品升级后在垂直领域形成数据飞轮先发优势，能够通过 SaaS（软件运营服务）模式、私有化部署、帮助企业搭建自有行业大模型等形式灵活接入。		
Whale 帷幄	Whale 帷幄基于 AGI，专为消费品牌群体和广大营销人而研发出来的 MarketingGPT Alivia ，只需通过人机对话的协作方式，就能轻松设计 Campaign；自动生成个性化图片和视频；根据个性化标签，精准一键投放；随时总结和分析数据。Alivia 从内容生产、管理到审核、发布，从数据分析到运营优化，整个营销运营工作的一整套闭环都能完成。	4 月 18 日	详情 1 ; 详情 2
创新奇智	面向制造领域的 AIGC 产品 奇智孔明 AlnnoGC 亮相。它是一个以工业预训练大模型为核心的生成式 AI 工程算法引擎，目前仍处于内测阶段。	4 月 13 日	详情
毫末智行	毫末智行官宣 DriveGPT，首个应用 GPT 模型和技术逻辑的自动驾驶算法模型，中文名 雪湖·海若 。参数量 1200 亿，模型终部署在云端。	4 月 11 日	详情

(3) 产品应用

AI 全家桶 Copilot

机构/个人名称	新闻动态	日期	链接
亚马逊	亚马逊旗下的云巨头 AWS 推出生成式 AI（AIGC）工具“全家桶”。AWS 产品覆盖 IaaS 实例、PaaS 平台、SaaS 软件全栈技术层。AWS 推出 生成式 AI 平台 Bedrock 服务 ，支持用户通过 API 访问亚马逊自己的 Titan（泰坦）大模型 （Titan 包含了两个大语言模型：一个是用于生成文本的 Titan text，一个是让网络搜索	4 月 14 日	详情 1 ; 详情 2 ; 详情 3

	个性化的 Titan Embeddings)；同时支持调用来自 AI21 Labs、Anthropic、Stability AI 等第三方的多样化模型。同时，Bedrock 一个核心特征是支持企业自定义基础模型，只需少量数据就完成模型定制化和微调。然后，AWS 推出了两款专门针对生成式 AI 优化的计算实例 EC2 Trn1n 实例和 EC2 Inf2。还推出一款 AI 编程伴侣 CodeWhisperer ，可辅助程序员编程，速度提升 57%；推出 AIGC 基础设施：基于自研芯片上线 EC2 Trn1n、Inf2 实例。		
阿里巴巴	阿里版的 Copilot 也要来了，阿里全系产品（钉钉、天猫精灵）也都将会接入“通义千问”大模型，包括文案、策划、邮件，几乎涵盖了所有衣食住行外加消费娱乐的场景。	4 月 11 日	详情 1 ; 详情 2

AI 编程助手

机构/个人名称	新闻动态	日期	链接
亚马逊	现已正式推出实时 AI 编程助手 Amazon CodeWhisperer ，包括 CodeWhisperer 个人套餐，所有开发人员均可免费使用，需使用电子邮件账户进行注册。对于企业用户，CodeWhisperer 提供专业套餐，增加了管理功能。	4 月 17 日	详情
谷歌	谷歌人工智能 Bard 学会写代码，修正代码里的 Bug，以及解释代码的用处，支持 C++、Java、Python 等 20 多种编程语言。同时，用户能将 Bard 生成的 Python 代码导出到 Google Colab 上，无需复制粘贴。Bard 还能够协助 Google Sheets 编写函数。谷歌也表示，Bard 仍处于早期试验阶段，用户必须始终仔细检查 Bard 的回复，在使用代码前仔细检查是否存在错误或漏洞。	4 月 21 日	详情

Replit	云开发平台 Replit 举办了自己的第一届开发者日活动。会上它宣布了自己的编码大模型 replit-code-v1-3b 。replit-code-v1-3b 拥有 27 亿参数，在 5250 亿代码 token 上进行了训练，可以识别 20 种语言。	4 月 27 日	详情
--------	--	----------	--------------------

AI 办公助手

机构/个人名称	新闻动态	日期	链接
askwise.ai	推出一个由 AI 驱动的个人知识助手 askwise 。能根据自己知识库定制 GPT 的 AI 应用。用户只需要上传 word、pdf、飞书文档、公众号网页链接等各种文档，就能生成个性化知识库，然后 AI 在你的知识库中进行搜索回答，知识库里没有的东西，它也能调用 GPT-3.5 的 API。目前功能：免费；文档数量无上限批量处理；双语提问，跨语言回答；提供答案来源索引等。	4 月 23 日	详情
石墨文档	4 月 18 日，石墨文档宣布接入百度文心一言，打造 AI 办公助手 石小墨 。目前，石墨文档在内测中已经研发出了快速创作的多个功能，用以辅助用户高效办公。AI 助手石小墨可以帮助用户撰写简历、故事、合同、博客等各种类型内容。	4 月 18 日	详情
腾讯	腾讯智能文档开启限量内测，新功能或对标 NotionAlo。发布的“智能文档”功能，将 All-in-one 的类 Notion 文档嵌入腾讯文档大家庭，用户可以在页面上乐高式拖拽块，制作高融合、无品类差异的应用文档。只要在文档空行处插入 “、” 或 “/” + 块名称的汉语拼音首字母即可插入相应的功能块。	4 月 18 日	详情
	3 月 30 日，腾讯正式发布 AI 智能创作助手“腾讯智影”，小程序已上线。“腾讯智影”是一款集成了 AI 创作能力的智能创作工具，提供虚拟数字人、文本配音、智能去水印、文章转视频、模板创作、在线视频剪辑等功	3 月 30 日	详情 ； 详情 2

	能。小程序已支持 AI 配音、视频审阅分享。		
字节-飞书	飞书公布了一 AI 智能助手 My AI ，My AI 可自动总结会议纪要，并根据会议纪要创建对应的待办事项。还可以根据文档、数据生成报告等内容，以及根据用户的编辑内容进一步续写。此外，My AI 还能够协助用户创建会议、查询案例参考、进行头脑风暴、规划项目进程等工作。但目前飞书方面尚未透露 My AI 的具体上线时间。	4 月 13 日	详情
印象笔记	印象笔记宣布 印象 AI 正式开放使用，现已落地赋能旗下全线产品，包括印象笔记、扫描宝、收藏家、印象时间、印象图记等。目前，用户已可下载最新版本体验，印象 AI 采用基于印象笔记自研的轻量化大语言模型大象 GPT 和国内外大语言模型相结合的混合模型方案。官方将在近期公布印象 AI 的商业化方案。	4 月 18 日	详情 1 ; 详情 2
阿里-钉钉	钉钉基于阿里通义千问大模型的“/”能力正式发布。根据发布的实时演示，用户只需在对话框里输入“/”，便可以唤起部分能力，完成内容生成、摘要提取、应用开发等诸多工作。	4 月 18 日	详情
百度	文心一言在百度内部全面应用在智能工作平台“ 如流 ”，已经为全体员工开通使用。可根据员工输入的提示词生成大纲、招聘 JD、新闻稿、宣传文案、邮件、调研报告等文案，也支持根据输入的文字生成匹配的图片，还提供代码推荐功能。目前百度如流的 AIGC 功能仅在公司内部上线，尚未对外公测。	4 月 18 日	详情
	百度文库开始内测“ AI 文档助手 ”，一句话生成文档的 AI 工具。创作文档仅需 15 秒，可以复制、导出文档。百度文库 AI 文档助手，接入的是百度文心一言，可提供活动方案、标书、研究报告、讲话稿等各种文档。	4 月 29 日	详情
北京大学深圳研究生院	研究团队日前开发了一款名为 ChatExcel 的 AI 办公辅助工具，可以通过文字聊天实现 Excel 的交互控制。ChatExcel 无需记函数、无需手动设置公式，只要在表格下的对话框内以文字输入要求，就能完成自运行，并可一键导出 Excel 表格。还可满足使用者持续交互、迭代更新的需求。ChatExcel 完全开放，测试者无需注册任何信	4 月 18 日	详情

	息、打开网页就能用、不限次数。 网址: https://chatexcel.com/		
金山办公	官宣将推出 WPS AI (暂名), 将用于旗下在线内容生成工具轻文档 (airPage), 并宣布金山办公全线产品即将嵌入 WPS AI。接入 “WPS AI” 后, AI 生成的内容可以直接嵌入到文档正文, 并能按照当前文档所能支持的格式进行实时渲染, 同时也支持多轮对话, 通过多次、连续自然语言的输入控制内容的生成。 “WPS AI” 提供起草、改写、总结、润色、翻译、续写等功能, 由国内合作伙伴 MiniMax 提供。	4 月 18 日	详情

浏览器/插件

机构/个人名称	新闻动态	日期	链接
OpenAI	Chrome 应用商店上线了插件 Monica, 它是 ChatGPT 在网页上的应用。Monica 可以聊天, 可以生成各种文案, 还可以阅读网页信息, 并且进行概括, 还能拓展出相关的问题作进一步解答。新用户有 30 次提问, 5 次阅读网页的机会。	4 月 17 日	详情
	官方推出了一种新模式 Default (GPT-3.5) with browsing。可以随时引用网络数据, 也可以直接读取指定网站中的文件了。目前仅部分用户可见。此前 ChatGPT “截止 2021 年数据” 的枷锁, 被彻底打破了。	4 月 28 日	详情
MLC.AI 社区	陈天奇团队发布 WebLLM , 无需服务器支持, 只需一个浏览器, 就能跑通大语言模型 (LLM), 基于 LLM 的类 ChatGPT 也能引进来, 而且不需要服务器支持、WebGPU 加速。主要流程是建立在 Apache TVM Unity 之上。 TensorIR 是生成优化程序的关键技术。该项目已在 GitHub 上开源。	4 月 19 日	详情
微软	4 月, 微软发布新版 Edge 浏览器, 其中包含了多个	4 月 10 日	详情

	新功能，包括图像生成、图像编辑、Drop、效率模式等。图像生成器（Image Creator）基于 OpenAI 的 DALL-E 模型。用户可以在不离开浏览器的情况下，找到社交、展示和其他用途的视觉图像，也允许用户创建不存在的图像。Drop 功能允许用户跨设备分享文件和笔记。	日	
--	---	---	--

AI 智能体

机构/个人名称	新闻动态	日期	链接
卡耐基梅隆大学	<p>该大学化学工程系的学者构建了一个基于大规模语言模型（LLMs）的自主科研智能体，实现了一个可以从自主设计、规划到执行复杂的科学实验的完整系统。它具有来自矢量数据库的长期记忆，可以阅读、理解复杂的科学文档，并在基于云的机器人实验室中进行化学研究。它会自己上网查文献，会精确控制液体处理仪器，还会解决需要同时使用多个硬件模块、集成不同数据源的复杂问题。还有自我纠错的功能，写出了以下论文。</p> <p>论文标题： <i>Emergent Autonomous Scientific Research Capabilities of Large Language Models</i></p> <p>论文链接： https://arxiv.org/abs/2304.05332</p>	4 月 13 日	详情 1 ; 详情 2
洛桑联邦理工学院、罗切斯特大学	<p>研究人员开发了 ChemCrow，一种 LLM 化学智能体，旨在完成有机合成、药物发现和材料设计等各种任务。通过集成 13 种专家设计的工具，并通过提示 LLM（GPT-4）提供有关任务和所需格式的具体说明来操作，ChemCrow 增强了 LLM 在化学方面的表现，并出现了新的功能。</p> <p>论文标题： <i>ChemCrow: Augmenting large-language models with chemistry tools</i></p> <p>论文链接： https://doi.org/10.48550/arXiv.2304.05376</p>	4 月 14 日	详情
OpenAI	<p>AutoGPT 是一个实验性的开源应用程序，由 GPT-4 驱动，可以自主实现用户设定的任何目标。它还具有互联网访问、长期和短期内存管理、用于文本生成的 GPT-4</p>	4 月 12 日	详情

	实例以及使用 GPT-3.5 进行文件存储和生成摘要等功能。AutoGPT 可用来分析市场并提出交易策略、提供客户服务、进行营销等其他需要持续更新的任务。AutoGPT 可以在 30 分钟内完成设置，就可以拥有自己的 AI，协助完成任务，提升工作效率。 GitHub 地址：https://github.com/torantulino/auto-gpt		
谷歌	谷歌近日公开全新网络完全套件 Google Cloud Security AI Workbench。该产品由名为 Sec-PaLM 的人工智能模型提供支持。Sec-PaLM 是谷歌 PaLM 模型的一个分支，针对安全用例对其进行微调。Google Cloud Security AI Workbench 包含了一系类的人工智能工具，可以利用 Sec-PaLM 来帮助用户查找、总结和应对安全问题。Google Cloud Security AI Workbench 目前还在测试之中，预计会在未来几个月内进行小范围测试。	4 月 25 日	详情
HyperWrite	创业公司 HyperWrite 发布了一个尚处实验阶段的 AI 智能体，这个智能体可以像人一样随意地浏览网页，甚至还能和网站进行互动。演示中，它能自行查找地址点披萨，填写邮编，付钱完成订单。	4 月 19 日	详情

5. 智能芯片

机构/个人名称	新闻动态	日期	链接
谷歌	谷歌设计了自己的定制芯片，称为 Tensor Processing Unit(TPU)，并将这些芯片应用于 90%以上的人工智能训练工作。目前，谷歌 TPU 已经进入第四代。对于规模相当的系统，TPU v4 可以提供比英伟达 A100 强 1.7 倍的性能，同时在能效上也能提高 1.9 倍。另外，谷歌超算速度还要比 Graphcore IPU Bow 快约 4.3 倍至 4.5 倍。谷歌利用自己定制的光开关将 4000 多个 TPU v4 芯	4月5日	详情 ; 详情2

	<p>片组合到一台超级计算机系统，用于训练人工智能模型。</p> <p>论文标题: <i>TPU v4: An Optically Reconfigurable Supercomputer for Machine Learning with Hardware Support for Embeddings</i></p> <p>论文链接: https://arxiv.org/abs/2304.01433</p>		
	<p>谷歌 Chrome 团队正式发布 WebGPU，用于在网络上进行高性能 3D 图形与数据并行计算。WebGPU 现已在 Beta 测试阶段的 Chrome 113 中默认启用。WebGPU 是一种新型 Web 图形 API，具有显著减少同等图形规模下 JavaScript 工作量、将机器学习模型的推理效率提升 3 倍以上等优势。WebGPU 的首个版本已经在 ChromeOS、macOS 和 Windows 上开放，对其他平台的支持将于今年晚些时候推出。</p>	4月6日	详情
微软	<p>微软正在开发人工智能 (AI) 芯片，为聊天机器人 ChatGPT 背后的关键技术提供动力。该计划在内部的代号为“雅典娜” (Athena)，2019 年已有 300 人共同研发，采用 5nm 技术。目前，芯片已经提供给了小部分微软的员工，以及 ChatGPT 开发商 OpenAI，主要用途为训练大语言模型和支持推理，是生成式人工智能 (AIGC) 背后的关键技术。</p>	4 月 19 日	详情 1; 详情 2
北京大学	<p>北大集成电路学院特聘研究员郑雨晴利用聚合物材料独特的可后修饰性，实现了聚合物导电、半导体性和介电材料的全光刻直接图案化，并成功制造出可随意拉伸、弯曲且性能稳定的弹性电路。该方法可在比拇指还小的面积 (0.238cm²) 上集成超过 1 万个弹性晶体管，创造了比此前报道提高 100 倍以上的器件阵列密度新纪录。这一技术突破了长期以来限制电子皮肤在实际工业生产应用中的关键瓶颈。</p> <p>论文标题: <i>Monolithic optical microlithography of high-density elastic circuits</i></p> <p>论文链接: https://www.science.org/doi/10.1126/science.abh3551</p>	4月5日	详情
	<p>北京大学人工智能研究院类脑智能芯片研究中心唐希源研究员课题组发布了一种最新研制并实现的增量型缩放</p>	4 月 20 日	详情

	<p>式模数转换器芯片，这款芯片在相似性能的增量型模数转换器中具有最高的输入带宽（150kHz）和最小的输入驱动要求（6.4pF/conv.），其能效达到 182.2dB FoM。</p> <p>论文名称： <i>A 150kHz-BW 15-ENOB Incremental Zoom ADC with Skipped Sampling and Single Buffer Embedded Noise-Shaping SAR Quantizer</i></p> <p>论文链接： https://ieeexplore.ieee.org/abstract/document/10067696</p>		
阿尔托大学	<p>研究人员以人类视网膜为灵感，发明了一种视觉传感器。其可以在视频的单一帧中识别移动的物体，并成功预测物体的移动轨迹，有望成为动态视觉传感、自动检测、工业过程控制、机器人引导和自动驾驶技术等领域的重要工具。该项技术的核心是一种对光产生电流反应的电子装置：光敏电阻阵列，可以有效地“记住”它们最近是否被暴露在光线之下。</p>	4 月 28 日	详情
香港大学	<p>助理教授向超实现光子芯片里程碑目标：在单个硅光芯片上集成激光与光频梳先进工艺。该研究的主要突破在于，首次实现了不需要光学封装、单片集成的氮化硅上激光器，并展现出在线宽、温度稳定性相比较传统半导体激光器的优势。</p> <p>论文标题： <i>Laser soliton microcombs heterogeneously integrated on silicon</i></p> <p>论文链接： https://www.science.org/doi/full/10.1126/science.abh2076</p>	4 月 20 日	详情
腾讯	<p>腾讯云发布面向大模型训练的新一代 HCC（High-Performance Computing Cluster）高性能计算集群。实测显示，该集群的算力性能较前代提升高达 3 倍。该集群采用最新一代腾讯云星星海自研服务器，搭载英伟达最新代次 H800 Tensor Core GPU。服务器之间采用业界最高的 3.2T 超高互联带宽，可以为大模型训练、自动驾驶、科学计算等提供高性能、高带宽和低延迟的集群算力。</p>	4 月 14 日	详情
	<p>腾讯自研芯片“沧海”已经广泛应用到实际环境，并斩获 8 项世界第一。在 4 月 15 日莫斯科国立大学举办的 MSU 硬件视频编码比赛中，编解码芯片“沧海”包揽了所参加的两个赛道 8 项评分的全部第一。腾讯自研的 AI</p>	4 月 15 日	详情

	推理芯片“紫霄”，采用自研存算架构和自研加速模块，可以提供高达 3 倍的计算加速性能和超过 45% 的整体成本节省，目前也已在内部业务中投用；高性能网络芯片“玄灵”，助力云计算场景实现虚拟化零损耗。		
之江实验室	<p>研究员虞绍良提出了新型光学芯片接口方案，将光学带宽提高 1 个数量级他提出的光学芯片新型通用接口方法，突破了因片上传统二维周期性结构维度缺失而带来的光场调控能力受限的瓶颈。该方法可达到超宽波段 90% 的光传递，实现了低损耗光学互连耦合方案，构建了波导集成的片上光镊系统。</p> <p>论文名称： <i>Free-form micro-optics enabling ultra-broadband low-loss off-chip coupling</i></p> <p>论文链接： https://onlinelibrary.wiley.com/doi/full/10.1002/lpor.202200025</p>	4 月 23 日	详情
字节跳动	字节跳动旗下火山引擎宣布自研 DPU (Data Processing Unit) 成功应用，已部署上万台 DPU 服务器。火山引擎宣布推出三款基于自研 DPU 的计算实例，分别基于 NVIDIA GPU、Intel CPU、AMD CPU，其中 NVIDIA GPU 计算实例相较上一代实现 3 倍性能提升。同时，火山引擎还推出了新版机器学习平台，上线模型效果对比等服务，平台支持单任务 GPU 万卡集群大模型训练、微秒级延迟网络；推出智能推荐-高速训练引擎，据称支持 100GB-1TB 超大模型高速训练。	4 月 18 日	详情
龙芯	龙芯近日发布了 3D5000 系列芯片，采用的是 Chiplet 技术，将两颗 3C5000 芯片封装在一起，实现了性能的 2 倍提升。两颗封装在一起后变成了 32 核，其性能已经不输 7nm 的 ARM 芯片，甚至不输 7nm 的 Zen2 芯片。	4 月 13 日	详情
蘑菇车联	蘑菇 AI 数字道路基站 2.0 (Mogo AI Station 2.0) 可以完成感知、计算、通讯，并下发红绿灯灯态、行人预警、路面状况预警、前车刹车预警等信息给车辆，帮助智能汽车安全实现路权博弈、提速、超车、自主变道等高难动作。	4月2日	详情
黑芝麻智能	4 月 7 日发布武当系列首款芯片 C1200 智能汽车跨域计算芯片平台。该芯片基于 7nm 制程，其使用支持锁	4月8日	详情

	步的车规级高性能 CPU 核 A78AE（性能高达 150KDMIPS）和车规级高性能 GPU 核 G78AE，提供强大的通用计算和通用渲染算力。该芯片内置成熟高性能的 Audio DSP 模块和每秒在线处理 1.5G 像素的新一代自研 NeuraIQ ISP 模块，提供 32KDIPMS 的行业最高 MCU 算力，能同时处理大于 12 路高清摄像头的输入，支持高速率的 MIPI。		
--	--	--	--

6. 智能机器人

机构/个人名称	新闻动态	日期	链接
麻省理工学院 MIT	科学家发明了新型机器手，可通过感知形状来识别它所抓的东西，准确率约为 85%。该实验装置 GelSight EndoFlex 采用了该大学的 GelSight 技术，该技术以前只用于机器人手的指尖垫。	4 月 4 日	详情
Levatas 、 波士顿动力	工程师 Santiago 宣布，成功在波士顿动力的 Spot 机械狗中，集成了 ChatGPT。根据演示，集成了 ChatGPT 的 Spot，能够回答关于它的电池电量水平、下一次任务的检查次数等问题，它也可以给出上一次任务的详细信息。除了语音回答外，在人工智能的加持下，Spot 还能够用点头或者摇头的方式来回应用户的提问。	4 月 27 日	详情 1 详情 2
阿里巴巴	阿里云工程师正在实验将千问大模型接入工业机器人，在钉钉对话框输入一句人类语言，即可远程指挥机器人工作。工程师可通过千问大模型自动生成代码指令完成机器人功能的开发与调试，甚至还能能为机器人创造一些全新的功能	4 月 27 日	详情
哈尔滨工业大学	团队构建出一种超分子胶体马达全新体系。该体系基于可控化学分子组装而成，并由数百个旋转生物分子马达（自然界中最小的分子马达）所驱动。该研究为设计游动纳米机器人仿生提供了一种新的指导思想；还为通过细胞的能量代谢主动调控，来进行疾病的精准诊疗提供了一种新方案。	4 月 5 日	详情
香港中文大学	张立教授团队提出一种构建可合并的微型机器人群体系统的策略，该系统利用铁磁流体液滴的分裂和聚合的特性来触发水平多模态和垂直抗重力集体的形成，并有多项功能。 论文标题: <i>Horizontal and Vertical Coalescent Microrobotic Collectives Using Ferrofluid Droplets</i> 论文链接: https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.202300521	4 月 6 日	详情
加拿大多伦多病童医院、多伦多	多伦多病童医院资深科学家黄熹博士和多伦多大学机械工程教授、机器人研究所所长孙钰教授提出了一种称为机械纳米手术的治疗肿瘤细胞新的方法，用精密磁控制来处理耐	4 月 14 日	详情

大学	化疗药物的 GBM。即使是对侵袭性、化疗耐药癌症也有效。通过使用纳米技术深入癌细胞，机械纳米手术就像“特洛伊木马”，可从内部破坏肿瘤细胞。		
西南交通大学	邓维礼/杨维清课题组构造出一种仿人类皮肤汗毛的微结构，这种仿生间歇微结构具有二级放大的功能，通过提高传感器的可压缩性，同时提高了传感器的灵敏度（461kPa-1）和响应范围（310kPa），其性能相较于常规的均一微结构高出好几倍。基于这种方法设计的柔性压阻式器件，具有灵敏度高、线性范围宽等特点。	4 月 6 日	详情
吉林大学	任露泉院士团队提出了一种新型液晶弹性体的 4D 打印策略，只需要在打印过程中对打印速度、挤出压力、打印高度和紫外光光强等多个参数进行设计，就能实现对液晶弹性体最终变形行为的控制。或在人工肌肉制造、柔性机器人等领域具备应用潜力。	4 月 10 日	详情
北京交通大学	第十三届“挑战杯”中国大学生创业计划竞赛全国决赛上，北京交通大学“Bixo——世界首创的双向线驱动下肢康复外骨骼机器人”团队获得了银奖。机器人可帮助下肢运动障碍的患者具备更好的行走能力，其助力效果较业内顶尖的两款哈佛大学样机提升近 60%，成本较市面产品降低 9 成；还更易穿戴和使用，为外骨骼在医疗领域的使用提供低成本高质量的解决方案。	4 月 7 日	详情
稚晖君团队	B 站野生钢铁侠（稚晖君）展示了双足机器人“哪吒”。哪吒的巧妙之处在于，一般被放置在大腿第一个关节的偏航角自由度转移到了最后一个关节。机器人便可以在保持运动学完全驱动可控的情况下，将末端的关节替换成一个自由转动的轮子。哪吒的脚就可以在“足”和“轮”之间来回切换，称之为“轮足二相腿”。	4 月 1 日	详情
美国 Sarcos	该公司的户外光伏板自主操纵 (O-AMPP)项目于 2021 年开始。Sarcos 机器人太阳能模块安装解决方案由配备 Guardian XM 智能机械手机机器人的自动工作车和优化太阳能电池板从交付到安装流程的自动送货车组成。该系统使用计算机视觉和人工智能软件传送、检测、提升和放置太阳能电池板。O-AMPP 系统可以将建造新太阳能发电场	4 月 17 日	详情

	所需的时间减少近 40%。公司计划于 2024 年将其安装太阳能电池板的自主机器人商业化。		
--	---	--	--

7. 语音识别

机构/个人名称	新闻动态	日期	链接
阿里巴巴	阿里大模型联合项目团队邀请脱口秀演员鸟鸟用手机录音 1 小时，做出一个“鸟鸟分鸟”模型。它不仅用宛如鸟鸟本尊的语气、语调跟人对话，而且用语用词方式与鸟鸟如出一辙。天猫精灵已接入“鸟鸟分鸟”模型，分别从孩子、老年人、工作党三类人群的视角，它能讲脱口秀段子，能陪用户唠嗑解闷，能检索日常问题，能回答天马行空的发散式问题，也能在察觉到用户心情不佳时送上安慰和建议。目前天猫精灵也在与第三方技术审计专家搭建一种 AI+ESG 的框架。	4月5日	详情

8. 信息检索与推荐

机构/个人名称	新闻动态	日期	链接
香港科技大学、北京大学、微软亚研等	研究者提出了一种高效利用多级用户意图的会话推荐模型 Atten-Mixer 。该模型可以与各种编码器集成，具备多级推理能力，在准确性和效率方面展现了优异的在线和离线性能。研究论文获得 WSDM2023 最佳论文荣誉提名。 论文标题: <i>Efficiently Leveraging Multi-level User Intent for Session-based Recommendation via Atten-Mixer Network</i> 论文链接: https://dl.acm.org/doi/abs/10.1145/3539597.3570445	4 月 13 日	详情
谷歌	谷歌正在开发新的搜索引擎，在现有功能的基础	4 月 17	详情

	增加人工智能。谷歌正在测试名为 Magi 项目的新功能，项目参与人员超过 160 人，新产品会试图预测用户需求。新的搜索引擎计划仍处于早期阶段，没有发布时间表。	日	
--	---	---	--

9. 安全与伦理

机构/个人名称	新闻动态	日期	链接
科技部	4 月 4 日，科技部发布关于公开征求对《科技伦理审查办法（试行）》意见的公告。此次《办法》提出，从事人工智能的单位，研究内容涉及科技伦理敏感领域的，应设立科技伦理（审查）委员会，并引入专家复核程序，部分算法模型、自动化决策系统需要专家复核。	4月5日	详情
澳大利亚	澳大利亚赫本郡市的市长发现，ChatGPT 在介绍 21 世纪初涉及澳大利亚储备银行一家子公司的海外贿赂丑闻时，错误地将他列为犯罪人员。他表示，如果 OpenAI 不纠正 ChatGPT 虚假指控他曾因贿赂入狱服刑的信息，他可能起诉 ChatGPT 的开发商 OpenAI，这将是针对人工智能机器人的第一起诽谤诉讼。	4月5日	详情
DeepMind	DeepMind 提出，政治哲学家罗尔斯的观点“无知之幕”可能有助于促进 AI 系统与人类价值观对齐过程中的公平性。 论文标题： <i>Using the Veil of Ignorance to align AI systems with principles of justice</i> 论文地址： https://www.pnas.org/doi/10.1073/pnas.2213709120	4 月 25 日	详情
三星	三星引入 ChatGPT 不到 20 天，就发生 3 起数据外泄事件。其中 2 次和半导体设备有关，1 次和内部会议有关。因为三星员工直接将企业机密信息以提问的方式输入到 ChatGPT 中，会导致相关内容进入学习数据库，从而可能泄漏给更多人。三星表示，已经告	4月2日	详情

	知员工谨慎使用 ChatGPT。如果后面仍发生类似事故，将考虑禁止在公司内网上使用 ChatGPT。		
--	--	--	--

10. AI 应用

(1) AI+医疗

机构/个人名称	新闻动态	日期	链接
复旦大学	<p>类脑智能科学与技术研究院冯建峰教授/程炜青年研究员团队联合国内外数十家医院，基于大规模脑影像数据（2000 多名被试），利用数据驱动的人工智能算法（疾病进展模型），推演出精神分裂症发生、发展的全维度疾病演进轨迹，精准定位出精神分裂症大脑萎缩的两个起源（语言区亚型和海马亚型）。</p> <p>论文标题: <i>Neuroimaging biomarkers define neurophysiological subtypes with distinct trajectories in schizophrenia</i></p> <p>论文链接: https://www.nature.com/articles/s44220-023-00024-0</p>	4月9日	详情

(2) AI+自然科学

机构/个人名称	新闻动态	日期	链接
伊利诺伊大学厄巴纳香槟分校	<p>赵惠民 (Huimin Zhao) 团队，开发了一种名为“启用对比学习的酶注释” (CLEAN, Contrastive Learning-enabled Enzyme Annotation) 的机器学习算法，可以根据酶的氨基酸序列预测酶的功能，即使这些酶尚未被研究或了解甚少。这款名为 CLEAN 的 AI 工具在</p>	4月4日	详情

	<p>准确性、可靠性和灵敏度方面超过了最先进的工具 (BLASTp)。</p> <p>论文名称: <i>Enzyme function prediction using contrastive learning</i></p> <p>论文链接: https://www.science.org/doi/10.1126/science.adf2465</p>		
华盛顿大学	<p>生物化学教授 David Baker 团队及其合作者, 将强化学习带入蛋白质设计, 开发了一款功能强大的新型蛋白质设计软件。实验结果显示, 由这种新方法合成的蛋白质能更有效地在小鼠体内产生有用抗体, 这一突破可能很快就会帮助科学家设计出更有效的疫苗。利用该方法, 团队设计出了上百个蛋白质结构。</p> <p>论文标题: <i>Top-down design of protein architectures with reinforcement learning</i></p> <p>论文链接: https://www.science.org/doi/10.1126/science.adf6591</p>	4 月 22 日	详情 1; 详情 2
清华大学	<p>交叉信息研究院曾坚阳团队通过小样本学习思想, 开发了能预测 T 细胞受体和抗原表位相互作用的深度学习模型 TEIM。模型首先在粗粒度的 TCR-epitope 序列结合信息上进行预训练, 然后在细粒度的残基相互作用信息上进行进一步微调。验证试验和分析表明, TEIM 模型具有优良的预测性能。</p> <p>论文标题: <i>Characterizing the interaction conformation between T cell receptors and epitopes with deep learning</i></p> <p>论文链接: https://www.nature.com/articles/s42256-023-00634-4</p>	4 月 10 日	详情
ByteDance Research、北京大学	<p>研究者将神经网络与扩散蒙特卡洛方法结合, 大幅提升神经网络方法在量子化学相关任务上的计算精度、效率以及体系规模。本工作表明, 基于神经网络的扩散蒙特卡洛方法在精度与效率上均优于变分蒙特卡洛方法。</p> <p>论文标题: <i>Towards the ground state of molecules via diffusion Monte Carlo on neural networks</i></p> <p>论文链接: https://www.nature.com/articles/s41467-023-37609-3</p>	4 月 12 日	详情

(3) 元宇宙/数字人

机构/个人名称	新闻动态	日期	链接
Rokid、钉钉	Rokid 携手钉钉共同打造了“钉钉数字文化墙”产品，基于 AR 创作工具“灵境”3D 点云、SLAM、3D 渲染等技术，把企业文化墙建设门槛降低，体验路径变短。通过钉钉微应用，只需普通手机、电脑或者 pad，就能在半小时内搭建一个全方位多模态交互的 AR 数字展厅。	4 月 20 日	详情
腾讯云智能	腾讯云智能小样本数智人生产平台首次对外发布，平台具有训练样本小、生产效率高、自动化生产等特性。只需要 3 分钟真人口播视频、100 句语音素材，平台便可通过音频、文本多模态数据输入，实时建模并生成高清人像，在 24 小时内制作出与真人近似的“数智人”。	4 月 25 日	详情
Snap	Snap 公司计划在实体商店推出增强现实（AR）镜子，可以让商店里的购物者立即看到穿上衣服的样子，而无需亲自试衣。这些 AR 镜子将于今年晚些时候入驻一些美国的耐克门店和男装连锁店 Men's Wearhouse。	4 月 23 日	详情

(4) AI+金融

机构/个人名称	新闻动态	日期	链接
嘉实基金	在 2023 年的基金一季报中，看好并专注投资人工智能的嘉实大科技研究总监、基金经理王贵重尝试用人工智能大模型写了其管理基金嘉实文体娱乐股票的一季报，引发了市场的关注。	4 月 22 日	详情
日本金融集团	日本金融集团加入了生成式 AI 的行列，开始采用 AI 驱动的聊天机器人来帮助完成报告和其他内部任务，减轻工作量。三菱日联金融集团计划在今年夏天推出	4 月 17 日	详情

	<p>ChatGPT 来起草申请批复文件和回复内部查询等工作，还考虑使用这项技术在线回答客户问题。</p> <p>三井住友金融集团已开始测试与微软日本合作开发的人工智能聊天机器人，还计划在今年秋季前后向所有员工推出这一工具。</p> <p>瑞穗金融集团正考虑与微软日本合作研发自己的聊天机器人，以保护敏感信息。</p>		
幻方量化	<p>对冲基金公司幻方宣布将成立一独立研究组织，专注于探索人工通用智能（AGI）。幻方 AI 已构建了深度学习训练平台“萤火二号”，它以“任务级分时共享”为核心理念，调度系统秒级响应。同时，平台配备强大的软件层支持：高性能算子库（hfai.nn）、分布式训练通讯框架（hfreduce）、大容量高带宽文件系统（3FS），让 AI 模型能自如拓展到多节点之上，进行大规模并行训练。</p>	4 月 17 日	详情

(5) AI+社会生活娱乐

机构/个人名称	新闻动态	日期	链接
清华大学、美国密西根大学	<p>科研人员合作研发出一套安全测试系统，为“AI 司机”量身定制出一套“驾考试题”。提出了密集强化学习方法（D2RL），通过识别和删除非安全关键状态、连接安全关键状态，并在编辑后的马尔科夫过程中训练神经网络，解决了“稀疏度灾难”。同时利用密集强化学习方法训练交通环境中的背景车辆，构建出一个由自动驾驶汽车和背景车辆组成的智能测试环境，从而实现了模拟环境替代实际道路环境。</p> <p>论文标题：Dense reinforcement learning for safety validation of autonomous vehicles</p> <p>论文链接：https://doi.org/10.1038/s41586-023-05732-2</p>	4月5日	详情
百度	4 月 3 日，基于文心交通大模型和生成式 AI 的底层支	4月3日	详情

	撑，百度地图 V18 版本更新，升级为人工智能新地图，面向实时交通实现刻画、感知、预测、调度的端到端极致优化，大幅提升效率。百度地图导航全新升级 3D 渲染引擎，同步支持手机和车机导航，实现全域毫秒级沉浸式出行体验。百度地图全新上线了领航数字人叶悠悠。		
浪潮	一个名叫“临小助”的数字社工助理诞生了，不仅能解答居民的提问，还能给社工做培训。“临小助”的这些能力的实现就是用到了浪潮大模型「源 1.0」的能力，该模型的参数规模 2457 亿，采用 5000GB 的中文数据集进行训练。	4 月 21 日	详情
上海人工智能实验室、商汤科技	通用视觉平台 OpenGVLab 的最新项目 Ask-Anything。它可以陪你看视频，边看边聊天。目前 Ask-Anything 提供了在线可玩的 demo，操作简单。只需要将视频“拖拽”或“上传”到页面左上角的指定位置即可。	4 月 25 日	详情
沃尔玛	美国最大零售商沃尔玛开始用名为“Pactum AI”的聊天机器人自动和供应商议价。它不仅让沃尔玛平均多省下 3% 的开销，而且每四个供应商中就有三个偏好与 AI 讨价还价，而非与人类沟通。沃尔玛只要告诉 Pactum AI 想要购买的产品规格、预算与付款条件等信息，Pactum AI 就能自动比较供货商的需求、价格趋势、以及其他竞品的价格等，再以这些信息去和供货商议价。	4 月 30 日	详情
中欧国际工商学院	近日发布《2022 年中国首席营销官（CMO）调查报告》。数据显示，在营销手段上，54% 的中国企业未使用新技术，在已使用新技术的企业中，17% 使用人工智能技术。以此推算，仅 7.8% 的企业采用人工智能技术营销。	4 月 12 日	详情

(完)

版权说明

本报告版权为智谱 AI 团队独家所有，拥有唯一著作权。本报告提供给订阅用户使用，仅限于用户内部使用。未获得智谱 AI 团队授权，任何人和单位不得以任何方式在任何媒体上（包括互联网）公开发布、复制，且不得以任何方式将研究报告的内容提供给其他单位或个人使用。如引用、刊发，需注明出处为“报告名称（智谱 AI 研究）”，且不得对本报告进行有悖原意的删节与修改。

本研究报告仅作为参考，智谱 AI 研究团队不保证分析得到的信息准确性和完整性，也不承担任何投资者因使用本产品与服务而产生的任何责任。

北京智谱华章科技有限公司（简称智谱 AI）主导研发了多语言千亿级超大规模预训练模型，构建了高精度通用知识图谱，并把两者有机融合打造了数据与知识双轮驱动的认知引擎。智谱 AI 提出全新 MaaS 的市场理念，打造了认知大模型平台、数字人和科技情报产品，应用单位包括：中国科协、北京市科委、华为、腾讯等 1000 余家企事业单位。智谱 AI 研究咨询部是专业从事科研信息情报研究咨询的机构，致力于为科技创新型客户提供重点领域科技及其人才情报、科研趋势研判、人才和机构技术尽调等专业服务。

欢迎交流合作，联系邮箱：miao.zhang@aminer.cn