

OPENSOURCE DAILY

开源日报

2024/1/31

第 006 期

开源中国技术社区

[OSChina.net](https://oschina.net)

引言

欢迎来到开源日报 —— 您探索和理解开源世界的日常指南。在这里，我们每天从宏观到微观，全面展现开源技术和文化的动态图景。开源日报致力于捕捉这个不断演进的领域的脉搏，为您提供开源技术如何塑造我们的世界和未来的深度洞察。

在宏观层面上，我们观察全球开源运动的大趋势，深入分析政策变化、市场趋势以及开源对全球技术生态的影响。

在中观层面上，我们聚焦于开源社区内部的动态，包括新兴技术的孕育、项目合作和行业标准的发展。我们深入探讨如何在实际操作中体现开源精神，为您带来关于行业最佳实践和标准的详尽分析。

而在微观层面上，开源日报密切关注具体的开源项目和个体贡献者。我们每天精选并分析热门的开源项目，探索它们的技术创新、应用实践及其对社会的影响。同时，我们讲述在这些项目背后辛勤工作的开发者们的故事，展现他们如何通过代码改变世界。

无论您是资深技术专家、行业观察者，还是对开源文化充满好奇的新手，开源日报都将为您提供独到的视角和深度的洞察。我们相信，每一行开源代码背后都有一个故事，每一个项目都蕴藏着无限可能。

加入我们的旅程，让我们共同揭开开源世界的奥秘，见证技术的力量如何每天以崭新的面貌展现在我们面前。

OSCHINA 编辑部

目录

1	数据概述	3
	(一) 开源社会数据	3
	(二) 今日开源环境指数	3
	(三) 开源话题情绪走势	4
	(四) 开源话题热词走势	4
	(五) 开源话题区域提及走势	5
	(六) 渠道传播 TOP	6
2	今日要点	7
	OpenSource Daily	7
3	今日观察	10
	1、社交观察	10
	1) 选择技术栈一定要选择开源的	10
	2) 技术创业非常难，就连 OpenAI 都不能说有护城河	10
	3) 深开鸿成功中标国家自然科学基金重点项目	11
	2、媒体观察	12
	1) 上海人工智能实验室发布开源开放评测体系“司南”	12
	2) 南大周志华团队 8 年力作！「学件」系统解决机器学习复用难题，「模型融合」涌现科研新范式	13
	3) IBM 着力宣传 Ceph，强调已在数据湖仓与生成式 AI 领域取得进展	13
4	事件点评	15
5	今日推荐	17

6	每日项目榜	19
	GitHub 榜单	19
7	用户观点	21
8	媒体观点	22
	附：中国基础软件开源产业主要参与者（持续更新）	23
	报告声明	26

1

数据概述

(一) 开源社会数据

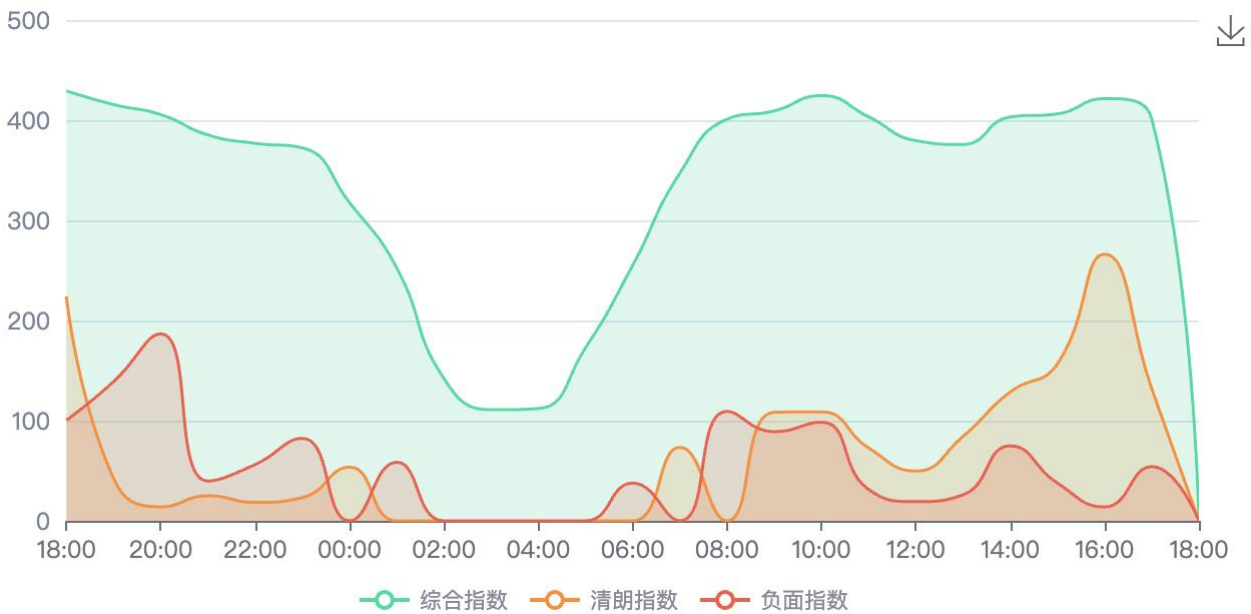
表 1：今日开源社会数据		
新闻平台	社交平台	视频平台
1621	1419	550
26.21%	22.96%	8.89%

备注：数据截止到 1 月 31 日

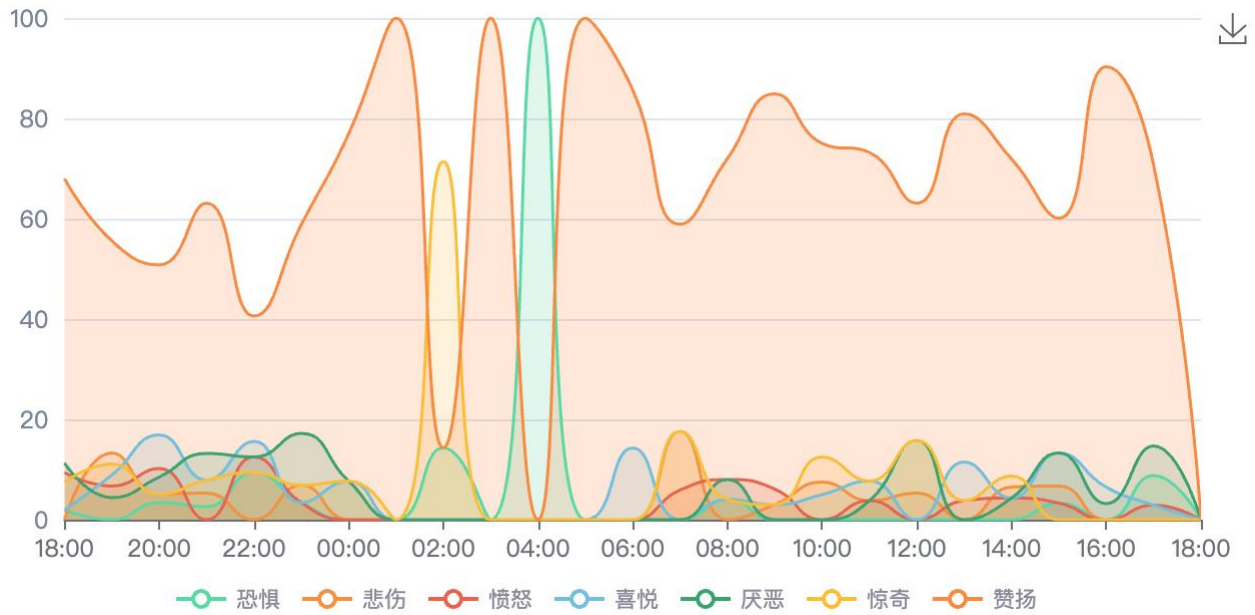
数据说明：

2024 年 01 月 30 日 18:00:00~2024 年 01 月 31 日 18:00:00，共监测到信息总量 6486 条，其中网页 1659 条（占比 25.58%），APP1518 条（占比 23.4%），微信 1148 条（占比 17.7%），视频 620 条（占比 9.56%），问答 515 条（占比 7.94%），论坛 409 条（占比 6.31%），微博 353 条（占比 5.44%），头条号 257 条（占比 3.96%），报刊 7 条（占比 0.11%）。

(二) 今日开源环境指数



(三) 开源话题情绪走势



数据说明：

分析时段内，赞扬 423 条（占比 66.81%）、喜悦 46 条（占比 7.27%）、厌恶 46 条（占比 7.27%）、惊奇 46 条（占比 7.27%）、愤怒 30 条（占比 4.74%）、悲伤 27 条（占比 4.27%）、恐惧 15 条（占比 2.37%）。

今日开源环境舆论中，喜悦和赞扬占比超 7 成，情绪一般

(四) 开源话题热词走势



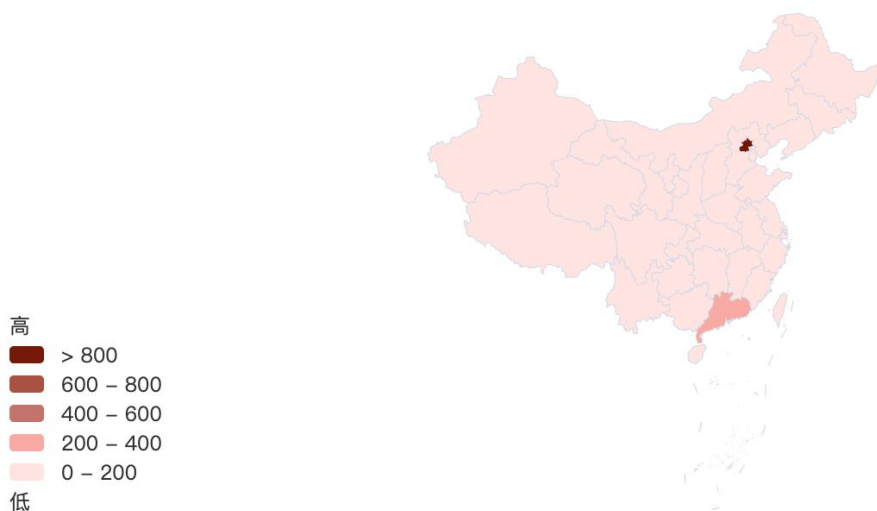
数据总结：

分析时段内，“模型”一词出现频率最高，达 2040 次，其次是“开源”，出现次数达 2020 次。

今日开源话题热词中，“开源”“GPT”“模型”“模态”等关键词在社交中热度持续上升

(五) 开源话题区域提及走势

地区	数量
北京	1010
广东	290
上海	194
安徽	51
浙江	49
河南	34
江苏	34
四川	30
福建	29
山东	27

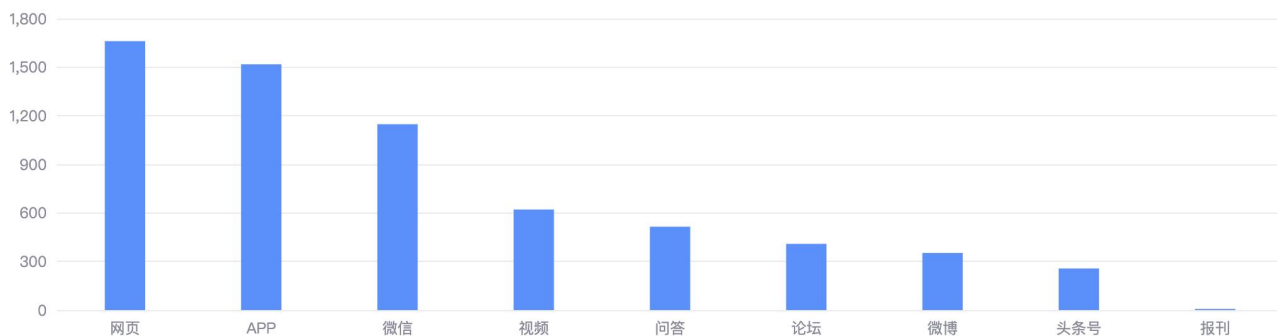


数据说明：

分析时段内，发布信息地域信息地域主要集中在北京 1010 条（占比 53.3%）、广东 290 条（占比 15.3%）、上海 194 条（占比 10.24%）、安徽 51 条（占比 2.69%）、浙江 49 条（占比 2.59%）。

今日北京、广东、上海地区对开源话题的兴趣度占比在 78% 以上。

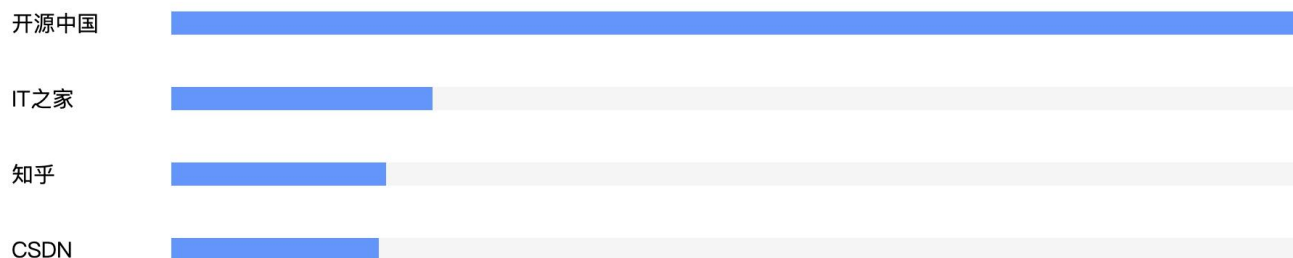
(六) 渠道传播 TOP



数据说明：

分析时段内，网页信息 1661 条（占比 25.6%）、APP 信息 1518 条（占比 23.39%）、微信信息 1148 条（占比 17.69%）、视频信息 621 条（占比 9.57%）、问答信息 515 条（占比 7.94%）、论坛信息 409 条（占比 6.3%）、微博信息 353 条（占比 5.44%）、头条号信息 257 条（占比 3.96%）、报刊信息 7 条（占比 0.11%）。

今日活跃度渠道评估：



2

今日要点

OpenSource Daily

Python HTTP 库 urllib3 支持直接在浏览器运行

Python HTTP 库 urllib3 最新版本 2.2.0 新增了一项重磅特性 —— 可直接在浏览器运行。

根据介绍,从 2.2.0 开始,urllib3 支持在 Pyodide 运行时中使用 JavaScript fetch API,或者在 fetch API 不可用时(例如当跨源隔离未激活时)返回 XMLHttpRequest。这意味着开发者可以使用该库直接从浏览器发出 HTTP 请求。

由于 urllib3 的 Emscripten 支持与 API 兼容,这意味着依赖于 urllib3 的库现在也可以在 Emscripten 和 Pyodide 环境中使用。

2.2.0 Latest

Compare

github-actions released this 11 hours ago · 2 commits to main since this release · 2.2.0 · 04df048

🖥️ urllib3 now works in the browser

🎉 This release adds experimental support for [using urllib3 in the browser with Pyodide!](#) 🎉

Thanks to Joe Marshall (@joemarshall) for contributing this feature. This change was possible thanks to work done in urllib3 v2.0 to detach our API from http.client. Please report all bugs to the [urllib3 issue tracker](#).

🚀 urllib3 is fundraising for HTTP/2 support

[urllib3 is raising ~\\$40,000 USD](#) to release HTTP/2 support and ensure long-term sustainable maintenance of the project after a sharp decline in financial support for 2023. If your company or organization uses Python and would benefit from HTTP/2 support in Requests, pip, cloud SDKs, and thousands of other projects [please consider contributing financially](#) to ensure HTTP/2 support is developed sustainably and maintained for the long-haul.

💬 编辑简评

urllib3 的最新版本引入了支持在浏览器中运行的特性,为开发者提供了更便捷的方式,这对于在浏览器中发出 HTTP 请求的应用场景具有积极意义。

热度评分



推荐星数



Opera 将推出全球首款非 WebKit 内核的 iOS 浏览器

昆仑万维旗下海外信息分发与元宇宙平台 Opera 近日宣布，将为欧洲的 iPhone 和 iPad 用户带来一款基于自研引擎的全新 AI 浏览器 “Opera One”。

Opera 计划于 3 月份首次亮相这款新浏览器，届时 DMA（欧洲数字市场法案）将生效。由于这些变化仅在欧盟实施，Opera 强烈鼓励苹果将这些增强功能推广至全球 iOS 用户。

为了遵守 DMA，苹果决定允许开发者提交非基于 WebKit 的浏览器，包括 Web 浏览器应用和在 iOS 应用中显示网页的应用内浏览器。Opera 表示，这些变化将使他们有机会为 iPhone 用户提供 Safari 之外的人工智能浏览器。

编辑简评

Opera 推出基于自研引擎的全新 AI 浏览器 “Opera One” 为欧洲 iOS 用户提供更多选择，尽管在全球范围内推广仍有待观察，但增强功能的引入为用户带来更多可能性。

热度评分



推荐关注



IntelliJ 负责人 Kirill Skrygan 成为 JetBrains 新任 CEO

JetBrains 公司正式对外宣布了最新人事任命：原负责领导 IntelliJ 部门的 Kirill Skrygan 将于 2024 年 2 月 1 日起担任首席执行官。

JetBrains 在公告写道：“在我们迈向人工智能发挥变革作用的新时代之际，公司已经建立了一个良好的底层基础，并且可以在此基础上继续推动公司向前发展。

作为新阶段的一部分，Kirill Skrygan 自 2010 年加入 JetBrains，并在过去几年里一直领导 IntelliJ 部门，他将于 2024 年 2 月 1 日担任首席执行官一职。”

JetBrains CEO Transition



JetBrains
January 30, 2024

The 1st of February 2024 will mark the 24th birthday of JetBrains. For more than two decades, JetBrains has gone from a company with a single product and one office, to a multinational entity with over 2000 people spread across 40 locations, and more than 30 products and services.

As we move towards a new era, in which AI is playing a transformative role, we have established a great base on which to build on and move this company forward. As part of this new phase, Kirill Skrygan, who has been with JetBrains since 2010 and for the past couple of years has been leading the IntelliJ Department will be stepping up as CEO effective February 1st 2024.

We remain strongly committed to our customers and partners and will continue to provide the products and services that they have come to expect and rely on us for. We'd like to specially mention Kotlin and The Kotlin Foundation in that it is strategic for JetBrains, and we will continue to invest into the language and ecosystem. – Kirill Skrygan

编辑简评

JetBrains 宣布新首席执行官任命，Kirill Skrygan 的经验和对公司的深刻了解为其领导角色提供了坚实基础，预示着 JetBrains 在人工智能时代的发展迎来新的里程碑。

热度评分



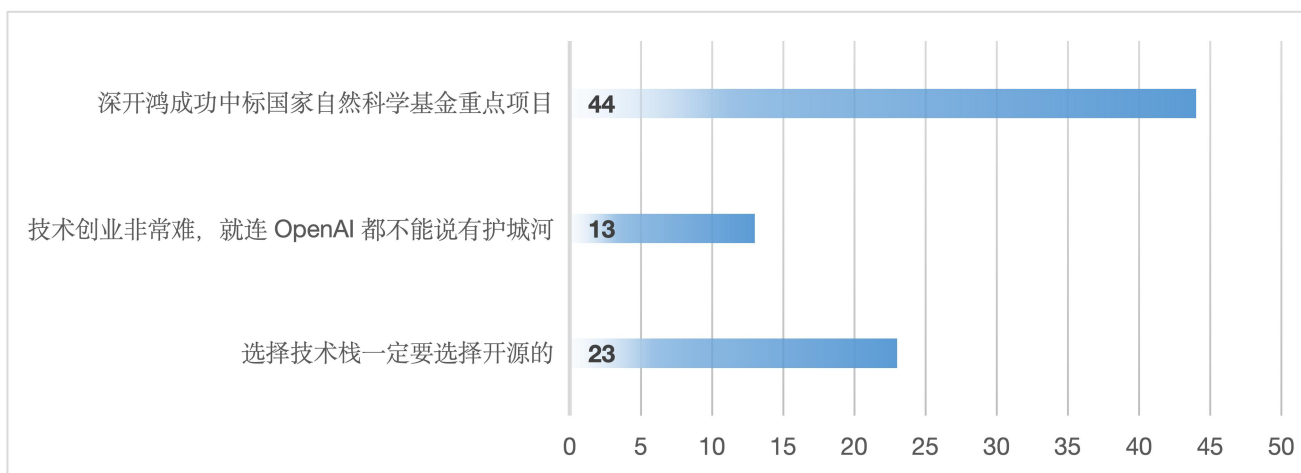
推荐关注



3

今日观察

1、社交观察



1) 选择技术栈一定要选择开源的

选择技术栈一定要选择开源的，附带加持是背后有大公司支持。只要开源，就基本上不会死，一种技术，不管流不流行，经过时间的洗礼，就会有或多或少的死忠分子，他们是开源生生不息的源泉。就怕专属某个公司的技术栈，说隔就隔，你哭都来不及。

- 微博 2gua

2) 技术创业非常难，就连 OpenAI 都不能说有护城河

技术创业非常难，就连 OpenAI 都不能说有护城河。

大语言模型 12 个月前是人人见而生畏的登月工程，现在借助开源产品，已经变成人人可上手。

周总说，原子弹变茶叶蛋。

这种特点，不能不注意。

- 微博 人马座 A 星的观星人

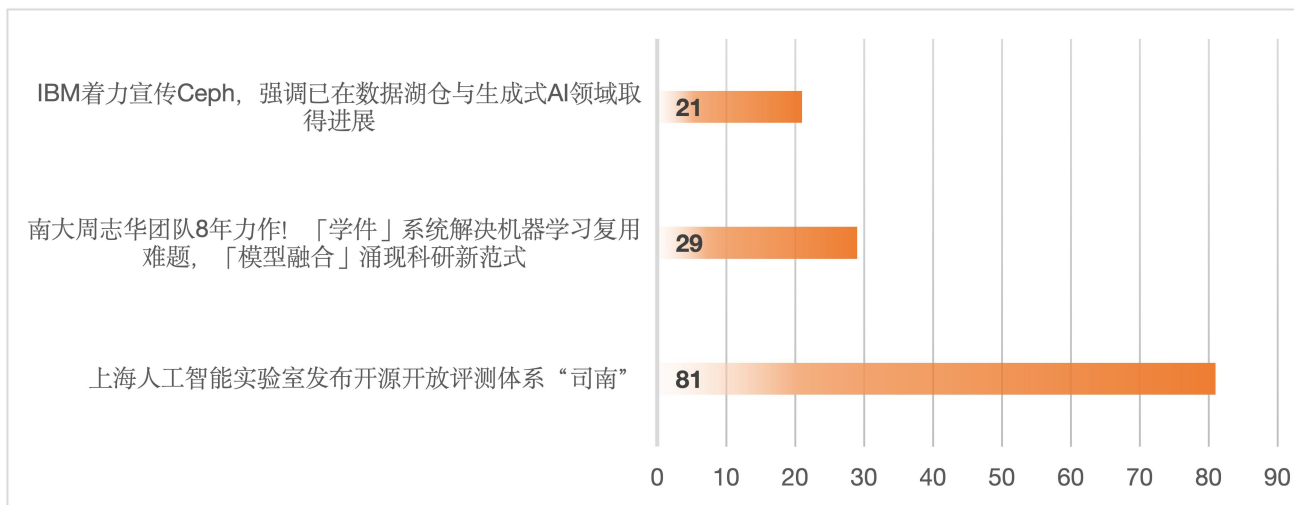
3) 深开鸿成功中标国家自然科学基金重点项目

深开鸿成功中标国家自然科学基金重点项目，将与北京大学、北京航空航天大学携手，共同开展开源软件可信安全研究。在 2023 年，该项目的申请通过率仅 16%，深开鸿凭借其卓越的技术实力和研究深度，成功获得国家和学术界的认可和支持！



- 微博 互联网科技报

2、媒体观察



1) 上海人工智能实验室发布开源开放评测体系“司南”

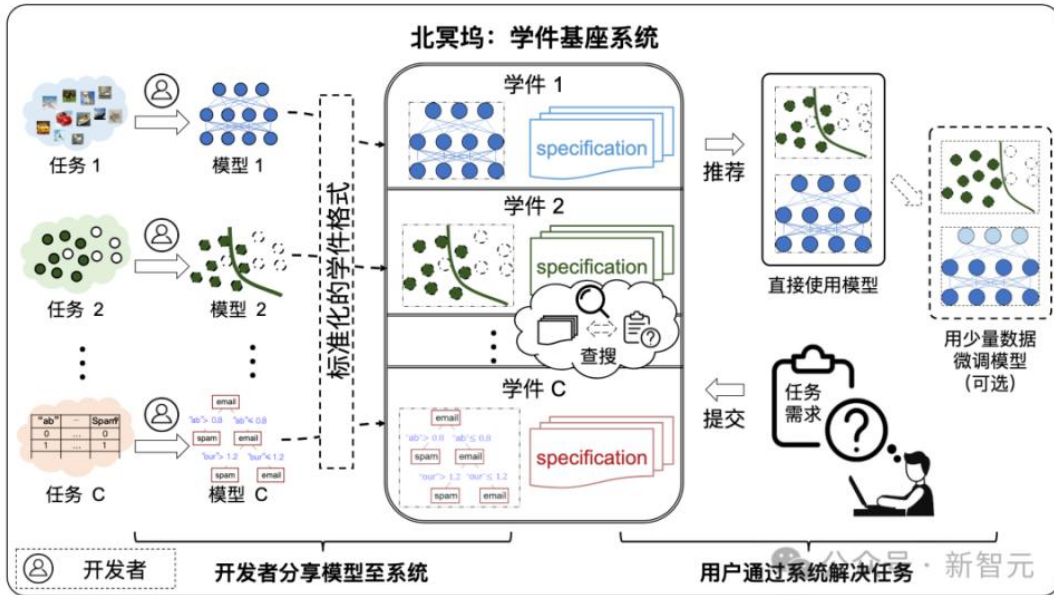
30日,上海人工智能实验室科学家团队正式发布大模型开源开放评测体系“司南”(OpenCompass2.0),可以为大语言模型、多模态模型等提供一站式评测服务。据介绍,“司南”全面量化模型在知识、语言、理解、推理和考试等五大能力维度的表现,评测榜单涉及的大语言模型和多模态大模型超过150个,客观中立地为大模型技术的创新提供坚实的技术支撑。截至目前,已有包括Meta、阿里巴巴、腾讯、百度等30余家国内外企业和科研机构采用“司南”助力开展技术研发。

Large Language Model				All	24-01	23-12
1	GPT-4-Turbo	61.8	OpenAI	6	MiniMax-abab5.5	50.7
2	GLM-4	58.9	ZhipuAI	7	InternLM2-Chat-20B	48.8
3	Qwen-Max	56.8	Alibaba	8	Yi-34B-Chat	47.1
4	Erniebot-4.0	54.7	Baidu Inc.	9	GPT-3.5-Turbo	46.8
5	Qwen-72B-Chat	51.9	Alibaba	10	Baichuan2-Turbo	46.4

- 新民晚报

2) 南大周志华团队 8 年力作！「学件」系统解决机器学习复用难题，「模型融合」涌现科研新范式

南京大学周志华教授团队推出的北冥坞，可以供研究者和用户像 HuggingFace 一样上传，使用机器学习模型。而且还创造性地通过基座系统，使不同的模型能够协同工作，应对复杂的任务，为全世界机器学习社区提供了一种全新的研究范式。



北冥坞不但可以让科研人员和用户像使用 HuggingFace 一样上传自己的模型，还可以基于基座系统，根据用户的需求进行模型匹配、协作融合，高效处理用户的学习任务。

- 新智元

3) IBM 着力宣传 Ceph，强调已在数据湖仓与生成式 AI 领域取得进展

自从 IBM 整合红帽旗下 Ceph 存储产品路线图已经过去了一年，如今这项技术已经在日益由 AI 主导的应用环境下取得进展。

IBM Storage Ceph 技术产品经理 Gerald Sternagl 在日前发布的博文表示，“这套具备自我修复与自我管理功能的平台，强调在行业标准硬件之上以规模化方式提供统一的文件、块和对象存储服务。统一存储平台将为客户提供重要的跨架构桥梁，能够将运行在独立文件或块存储上的遗留应用程序同对象存储共同

囊括在同一设备之内。”

博文解释称，“在数据管理领域，软件定义存储已经成为一股变革性的力量。与传统存储阵列相比，其拥有诸多优势，包括极高的灵活性与可扩展性，特别适合处理生成式 AI 等现代用例。”

Sternagl 对于 IBM 的传统存储阵列硬件（例如大型机 DS8000 以及 x86 服务器 FlashSystem 阵列等）持批评态度。且在被蓝色巨人收购之前，他已经作为资深专家为红帽工作了十余年。

在他看来，“Ceph 针对大型单站点与多站点部署进行了优化，能够高效扩展以支持数百 PB 的数据和数亿个存储对象，从而为传统及新兴生成式 AI 等工作负载提供至关重要的支持。” Ceph 能够适应数据湖仓、AI/机器学习开源框架以及其他“较为传统的工作负载，包括红帽 OpenShift 以及 RedHat OpenStack 上的 MySQL 与 MongoDB。”

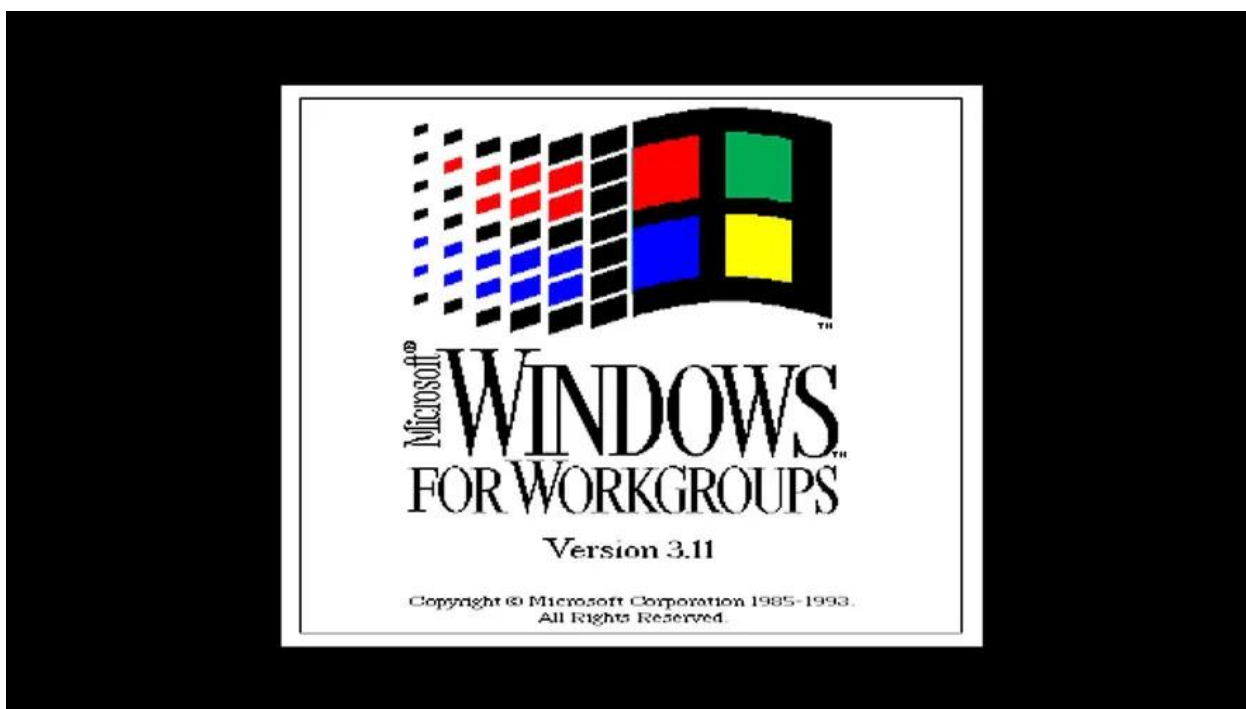
- 至顶网

4

事件点评

德国铁路公司招聘熟悉 MS-DOS 和 Windows 3.11 的 IT 管理员

推特用户 @konkretor 发帖介绍了一则乍看让人感觉时空错乱的招聘消息：一家德国铁路公司正在寻求熟悉 Windows 3.11 系统的 IT 管理员——有 MS-DOS 使用经验更佳。据称该岗位负责“德国几乎所有铁路的显示面板”。



招聘信息写道，该岗位的工作内容是维护对铁路运营仍然至关重要的旧系统，以及更新驱动。它们运行于高速和区域列车上的驾驶室显示面板，用于向驾驶员实时展示最重要的技术数据。

事件点评

微软在 1992 年推出了 Windows 3.1，至今已有 30 多年历史。德国的工业和科技整体上起步较早，所以像

铁路这样的关键基础设施依靠着“上古时代”的操作系统和硬件并不罕见。更何况这个岗位负责的显示面板只是用于展示信息数据，不涉及到关键功能——“如果没有坏到不能用，就用不着更换”。

自身影响力	辐射影响力
★★★★☆	★★★★☆

分析了 1.5 亿行代码发现：AI 编程助手降低代码质量

根据 GitHub 在 2022 年发表的一项研究，使用 GitHub Copilot 的开发者完成任务的速度更快，比不使用 GitHub Copilot 的开发者快 55%。但 GitClear 的最新报告则指出 AI 编程助手降低了代码质量和可维护性。报告指出，GitHub Copilot 与向代码库推送“错误代码”强相关，AI 编程助手倾向于不复用代码，倾向于更多拷贝粘贴代码，增加了未来的维护负担。报告认为 GitHub Copilot 是有前途的工具，但目前还有很多问题。

事件点评

借助 AI 提供商，您可以将代码生成速度提高 50%（即使是您不理解或无法编写的代码），但代价是代码的质量和可持续性不断下降。我们要追求的，到底是质量还是速度？

自身影响力	辐射影响力
★★★★☆	★★★★☆

5

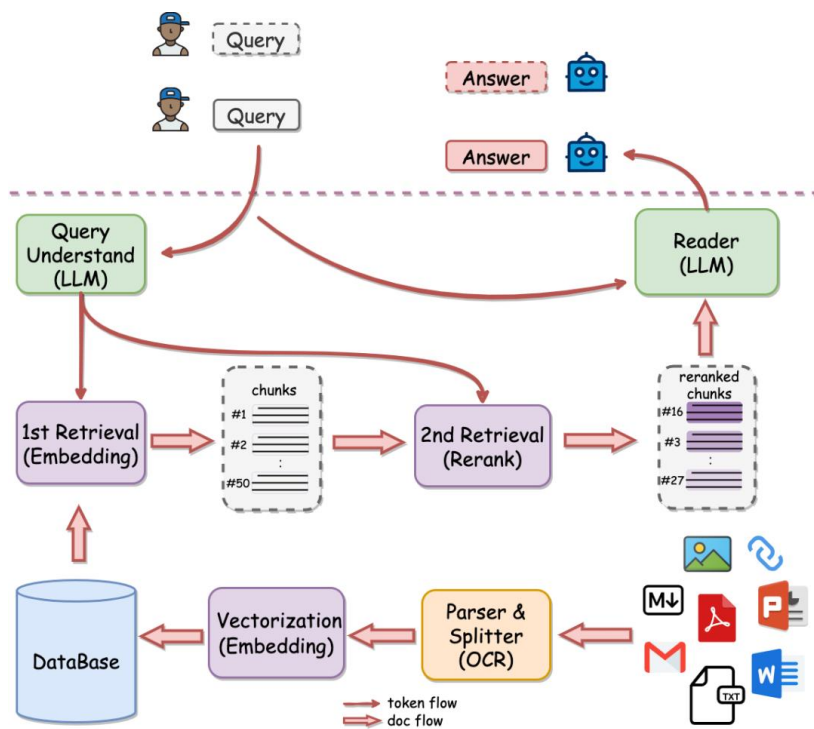
今日推荐

netease-youdao / QAnything

项目介绍:

QAnything (Question and Answer based on Anything) 是致力于支持任意格式文件或数据库的本地知识库问答系统，可断网安装使用。

您的任何格式的本地文件都可以往里扔，即可获得准确、快速、靠谱的问答体验。目前已支持格式: PDF(pdf), Word(docx), PPT(pptx), XLS(xlsx), Markdown(md), 电子邮件(eml), TXT(txt), 图片(jpg, jpeg, png), CSV(csv), 网页链接(html)。



Stars	Forks	Watching
3.6k	290	37

推荐理由：

QAnything 的主要原理是基于检索增强的生成（Retrieval Augmented Generation，简称 RAG），RAG 是目前很火的一种技术框架，它能够利用检索外部内容的方式增强大语言模型的准确度、专业能力和个性化等各方面的性能。

目前，通用大模型能力很强，但无法使用用户的私有数据，且存在胡乱编造的幻觉问题。QAnything 作为有道自研的 RAG 引擎，则结合了用户私有数据和大模型的优势——用户的任何内容，以任意的形式存在，比如各种格式的文档，音频，数据库等，都可以在 QAnything 的支持下，变成可以针对其内容进行问答的使用方式，通过这个技术框架用户可以很方便地搭建自己的智能知识助手。

6

每日项目榜

GitHub 榜单

Rank	Language	Stars	Forks	Stars Today
01	leptonai/search_with_lepton			
	Building a quick conversation-based search demo with Lepton AI.			
	TypeScript	3325	325	1348
02	rasbt/LLMs-from-scratch			
	Implementing a ChatGPT-like LLM from scratch, step by step			
	Jupyter Notebook	5211	284	806
03	vikhyat/moondream			
	tiny vision language model			
	Python	1210	89	120
04	KaisenAmin/c_std			
	Implementation of C++ standard libraries in C			
	C	889	49	233
05	InkboxSoftware/excelCPU			
	16-bit CPU for Excel, and related files			

	Python	1312	105	493
06	01-ai/Yi			
	A series of large language models trained from scratch by developers @01-ai			
	Python	5930	364	47
07	danielmiessler/fabric			
	fabric is an open-source framework for augmenting humans using AI.			
	Python	524	67	98
08	TaskingAI/TaskingAI			
	The open source platform for AI-native application development.			
	Python	1086	123	311
09	Fplasma-umass/scalene			
	Scalene: a high-performance, high-precision CPU, GPU, and memory profiler for Python with AI-powered optimization proposals			
	JavaScript	10815	377	30
10	saagarjha/Ensemble			
	Cast Mac windows to visionOS			
	Swift	414	15	73

7

用户观点

开源工具搜索toollm	
选开源的也要评估过自己能不能hold住	
期待他们在开源软件安全领域取得更多突破性成果，真的是太不错了	
今日AI日报	
转发	
啊，主要是激情兄弟自己的这个做的不行，虽然不喜欢粗粮，但是粗粮的app真的挺好的，不过洗衣机冰箱真的没啥好控制的	
虽然接触技术的时间不长	

1、IntelliJ 负责人 Kirill Skrygan 成为 JetBrains 新任 CEO

观点 1：希望 JetBrains 从卖内存条和 CPU 的公司回到卖软件。

观点 2：恭喜，只不过太卡了，MacBook M1Pro 32G 内存都卡。。。

2、微软组建新团队，帮助用 Rust 重写核心 Windows 库

观点 1：用 Rust 替换 C# ？这是要抛弃“亲儿子”了吗？

观点 2：rust 值得学

3、德国铁路公司招聘熟悉 MS-DOS 和 Windows 3.11 的 IT 管理员

观点：“高明的人往往都是用着简单实用的东西。稳定压倒一切，你猜英国核潜艇用的是什么系统？win95

8

媒体观点

1、钛媒体

观点：李彦宏替马化腾试错？

围绕 AI 原生应用，二人的态度并不一致，两家公司做大模型的思路、布局路径也不一样。

比起百度，腾讯对大模型在应用侧上的判断略加保守，而他们在认知上的这种不同，映射到具体的业务侧，也呈现出来了不同的战略姿态。

2、第一财经

观点：国内大模型与 GPT-4 较高下：语言知识能力接近，复杂推理仍是短板

关于评测，排名可能并不是最需要关注的，在榜单上一时的排名高或低并不能真正反映大模型的能力，评测的真正价值是帮助机构和企业发现自家大模型进一步需要努力的方向。

3、北京商报

观点：升级大模型至 3.5 版本，科大讯飞“秀肌肉”

“大模型发展，应用才是硬道理。”刘庆峰强调。最新数据显示，自 2023 年 5 月讯飞星火正式发布起，有超过 60 万讯飞智能办公本、智能录音笔用户累计调用语篇规整、会议纪要、自动写稿等能力 650 万次。“大模型竞争一年，现在已经到了落地应用的 PK 阶段。这一阶段的竞争将持续很长时间。”

附：中国基础软件开源产业主要参与者（持续更新）

基础软件开源项目

AI 框架



大模型



数据库



中间件



操作系统



开源组织

开源社区评估机构



开源技术论坛



开源代码托管平台



开源基金会



开源产业联盟



报告声明

分析法说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。

社区声明

本报告团队使用合法合规的信息，独立、客观地出具本报告，并对本报告的内容和观点负责。所有编辑人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的意见或观点有直接或间接的联系。

免责声明

《开源日报》是开源中国技术社区编辑人员对开源内容相关信息和数据的综合量化评估分析，其性质是为单位或个人研究与观察开源领域发展提供辅助，并不构成最终决策建议。材料结论、数据的完整性、适用性，不予保证，请单位和个人斟酌参考。

版权声明

《开源日报》为开源中国技术社区所作，未经开源中国技术社区的审核、确认及书面授权，不得以任何方式在任何媒体上（包括互联网）公开引用本报告的数据和观点，不得以任何方式将本报告的内容提供给其他单位或个人，否则引起的一切法律后果自行承担，同时开源中国技术社区团队亦认为其行为侵犯了开源中国社区团队的著作权，开源中国技术社区团队有权依法追究其法律责任。