

OPENSOURCE DAILY

# 开源日报

2024/2/2

第 008 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。

## 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 .....	11
	1、社交观察 .....	11
	1) 北京浮生记 GPT .....	11
	2) 12036 抢票助手开源, 星标 12.7K ! .....	12
	3) ChatLLM.cpp: 纯 C++实现的库 .....	13
	2、媒体观察 .....	14
	1) Red Hat 合作伙伴计划变更, 旨在实现“100%合作伙伴主导的驱动力” .....	14
	2) 推动中国开源软硬件发展的经验与建议 .....	14
	3) 快手前技术副总裁王仲远加入智源研究院 接替黄铁军担任院长 .....	15
4	事件点评 .....	16
5	今日推荐 .....	19
6	每日项目榜 .....	20

GitHub 榜单 .....	20
7 用户观点 .....	22
8 媒体观点 .....	23
附：中国基础软件开源产业主要参与者（持续更新） .....	24
报告声明 .....	27

# 1

## 数据概述

### (一) 开源社会数据

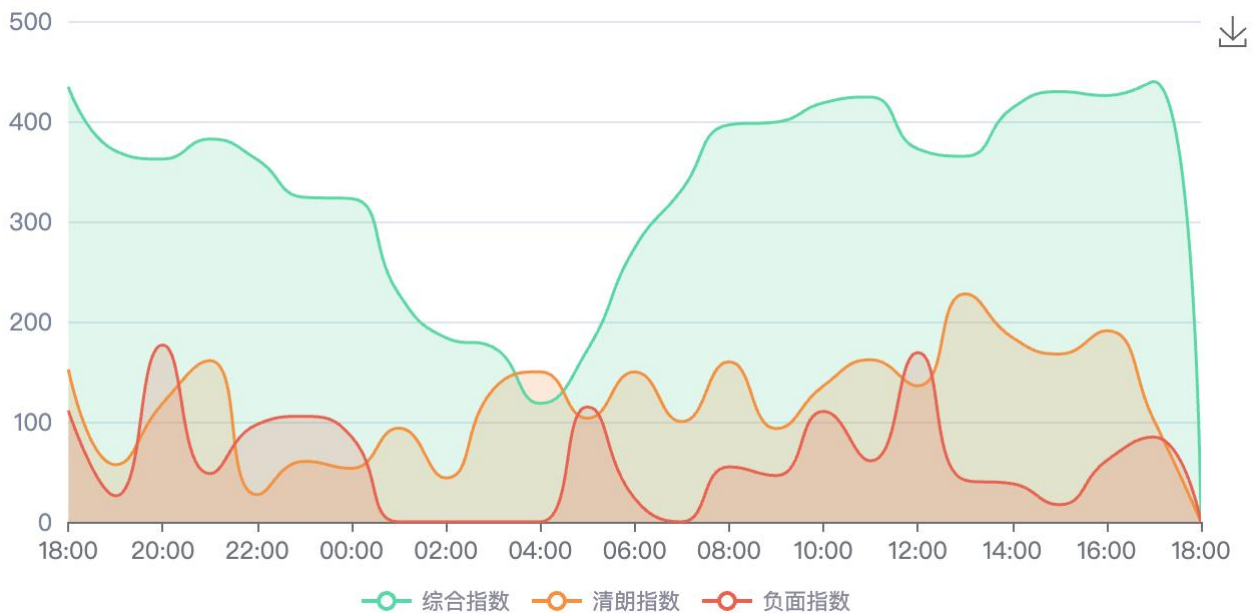
表 1：今日开源社会数据		
新闻平台	社交平台	视频平台
1600	1636	527
26.43%	27.02%	8.71%

备注：数据截止到 2 月 2 日

#### 数据说明：

2024 年 02 月 01 日 18:00:00~2024 年 02 月 02 日 18:00:00，共监测到信息总量 6053 条，其中网页 1600 条（占比 26.43%），APP1233 条（占比 20.37%），微信 1203 条（占比 19.87%），视频 527 条（占比 8.71%），问答 469 条（占比 7.75%），微博 433 条（占比 7.15%），论坛 313 条（占比 5.17%），头条号 260 条（占比 4.3%），报刊 15 条（占比 0.25%）。

### (二) 今日开源环境指数





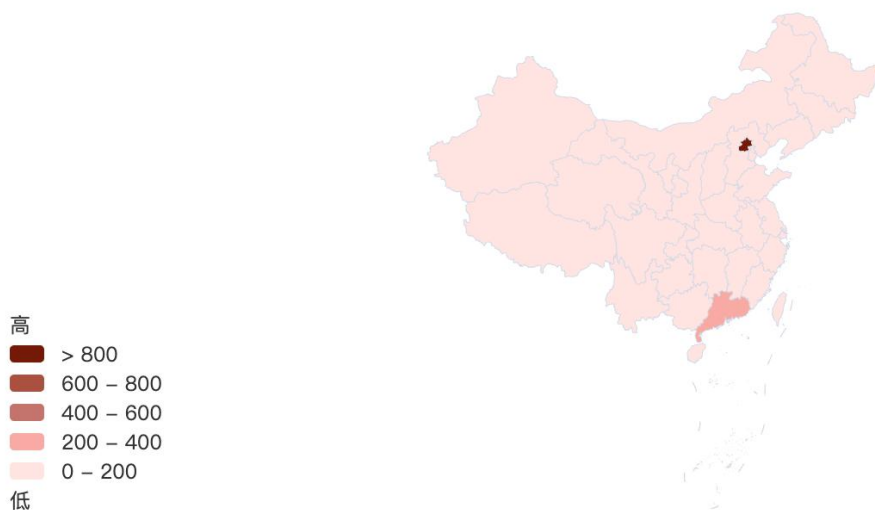
数据总结:

分析时段内,“开源”一词出现频率最高,达 2131 次,其次是“模型”,出现次数达 1759 次。

今日开源话题热词中,“开源”“AI”“模型”“数据库”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	879
广东	289
上海	127
江苏	37
河南	35
湖北	33
浙江	32
山东	31
山西	29
内蒙古	25

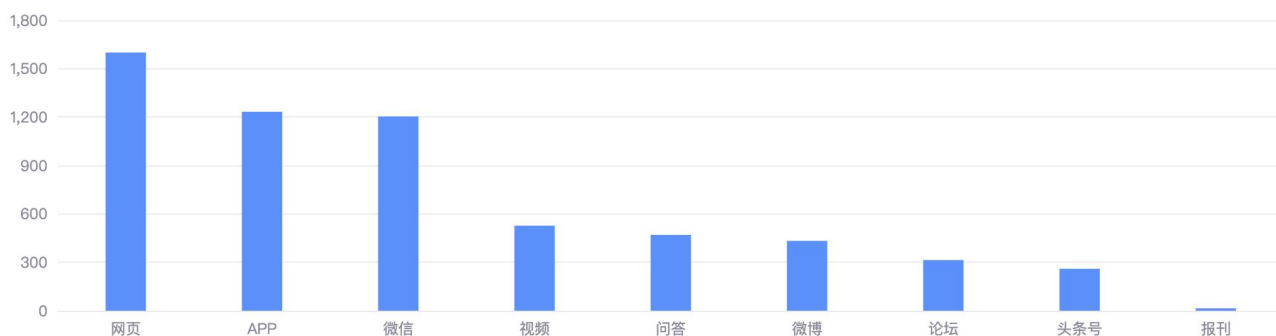


数据说明:

分析时段内,发布信息地域信息地域主要集中在北京 879 条(占比 51.68%)、广东 289 条(占比 16.99%)、上海 127 条(占比 7.47%)、江苏 37 条(占比 2.18%)、河南 35 条(占比 2.06%)。

今日北京、广东、上海地区对开源话题的兴趣度占比在 75% 以上。

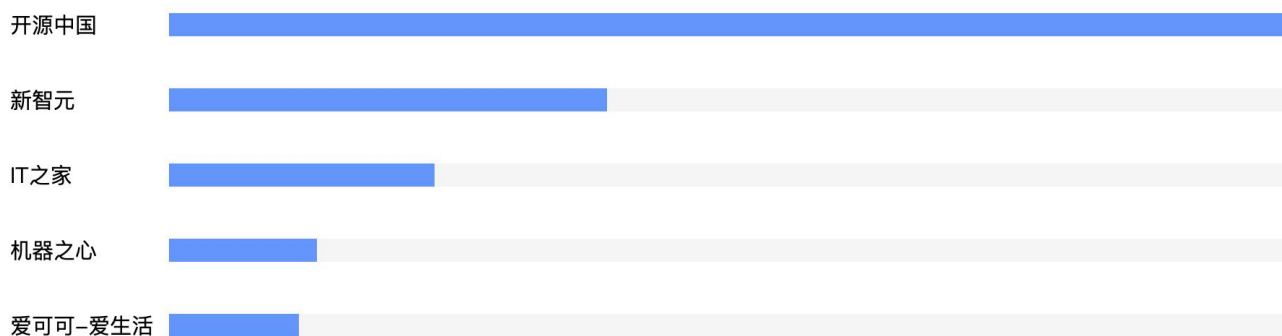
## (六) 渠道传播 TOP



### 数据说明：

分析时段内，网页信息 1600 条（占比 26.42%）、APP 信息 1233 条（占比 20.36%）、微信信息 1204 条（占比 19.88%）、视频信息 527 条（占比 8.7%）、问答信息 470 条（占比 7.77%）、微博信息 433 条（占比 7.15%）、论坛信息 314 条（占比 5.18%）、头条号信息 260 条（占比 4.29%）、报刊信息 15 条（占比 0.25%）。

### 今日活跃度渠道评估：





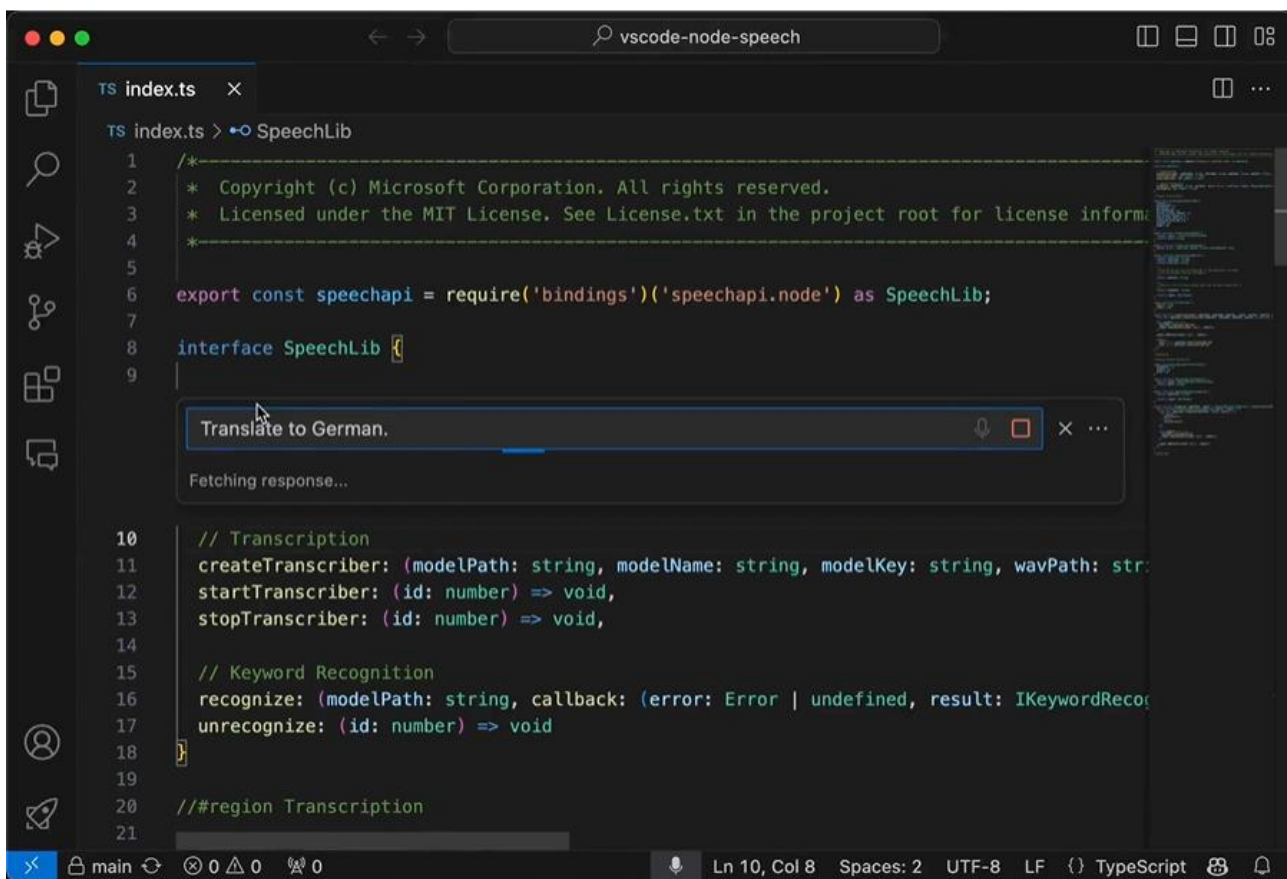
# 2

## 今日要点

### OpenSource Daily

#### Visual Studio Code 1.86 发布

新版本支持独立调整每个窗口的缩放级别、添加 Hey Code 语音命令——支持在差异编辑器中快速查看多个文件的差异、支持使用语音命令启动聊天会话。



#### 编辑简评

新版本通过支持独立调整每个窗口的缩放级别，开发者能够更方便地适应不同工作区的需求。引入 Hey Code

语音命令和支持通过语音启动聊天会话则进一步提高了开发效率，为使用 Visual Studio Code 的开发者带来更多便捷的操作方式。这一系列更新旨在提升开发者的工作流程和用户体验，是一次令人期待的升级。

热度评分	推荐星数
★★★★☆	★★★★☆

### AI2 开源新 LLM，重新定义 open AI

艾伦人工智能研究所（Allen Institute for AI，简称 AI2）宣布推出一个名为 OLMo 7B 的新大语言模型，并开源发布了预训练数据和训练代码。OLMo 7B 被描述为 “一个真正开放的、最先进的大型语言模型”。

AI2 是由已故的微软公司联合创始人保罗 - 艾伦（Paul Allen）于 2014 年在西雅图创办的一家非营利组织。该组织表示，他们开源 OLMo 模型旨在让 AI 研究社区充分了解最先进的大语言模型（LLM），从而推动自然语言处理的发展，并以科学的方式正视现有 LLM 存在的问题。



“通过向公众完全开放 OLMo 及其训练数据, AI2 在协作构建世界上最好的开放语言模型方面迈出了一大步。”

 编辑简评

AI2 发布的 GenAI 语言模型系列, 尤其是 Dolma 模型和 OLMo 数据集, 标志着在开放领域取得的显著进展。

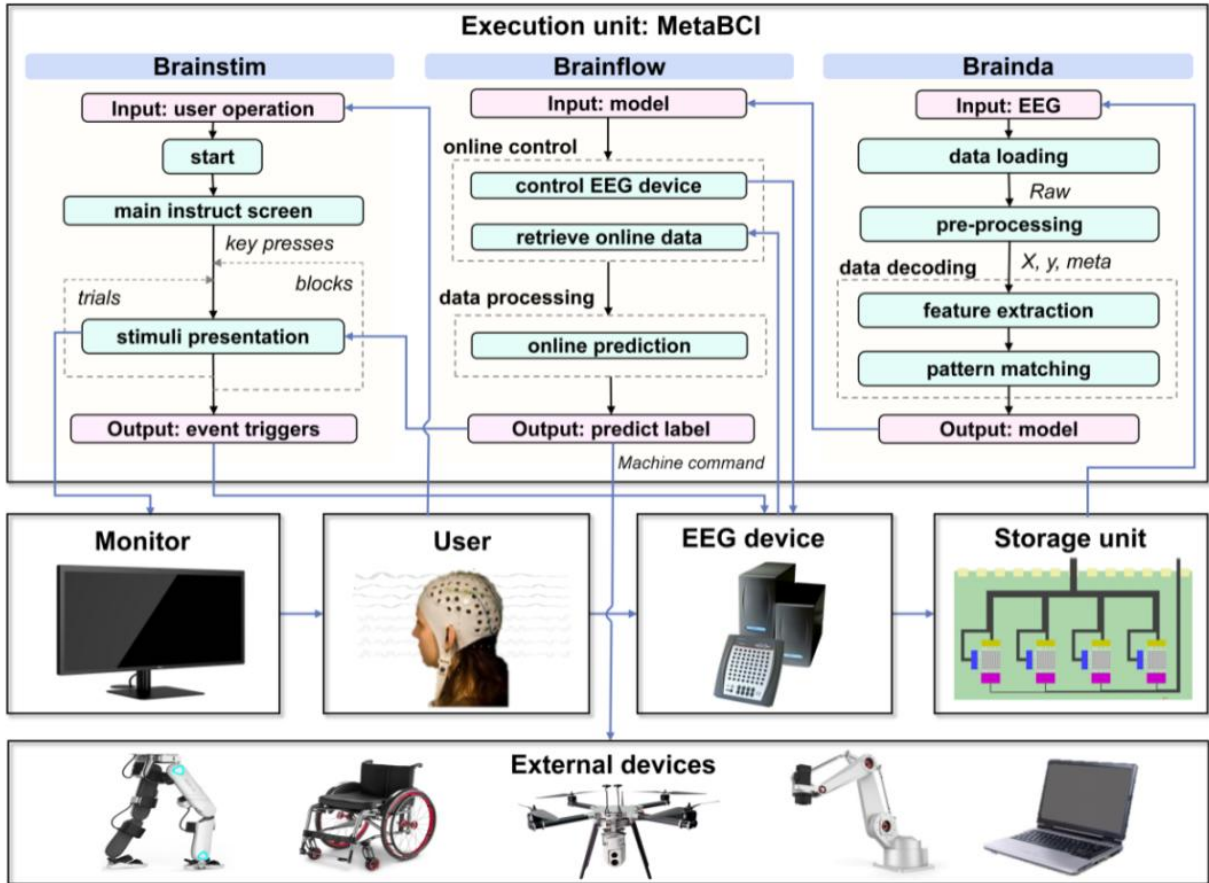
通过提供开放、无限制使用的许可方式, 以及强调高级科学研究, AI2 积极推动语言模型领域的开放性、透明度和创新, 为研究者和开发者提供了强大的工具和资源。

热度评分	推荐关注
★★★★★	★★★★★

天津大学发布中国首个脑机接口开源软件平台 MetaBCI 研究成果

近日, 天津大学神经工程团队在数学与计算生物学领域 SCI 一区 TOP 期刊《Computers in Biology and Medicine》上发表论文, 介绍了中国首个脑机接口开源软件平台 MetaBCI 的技术架构与实现方法。”

MetaBCI 由离线分析模块 Brainda、刺激呈现模块 Brainstim 和在线数据流模块 Brainflow 三大部分构成, 提供了面向 BCI 软件层面全链条开发的解决方案。面向离线分析需求, Brainda 统一了现有公开数据集接口, 优化了脑电数据读取、处理流程, 复现多种主要 BCI 数据分析及解码算法, 以此提高研究者的算法开发效率; 面向刺激呈现需求, Brainstim 提供了简洁高效的范式设计模块, 可快速创建脑机接口范式刺激界面; 面向在线开发需求, Brainflow 利用双线程、双进程编程方法实现了实时高速的数据读取、数据处理、结果反馈等功能, 帮助开发者轻松搭建脑机接口在线实验系统。



编辑简评

天津大学神经工程团队推出中国首个脑机接口开源软件平台 MetaBCI，该平台结构完备，涵盖了离线分析、刺激呈现和在线数据流等三大关键模块，为 BCI 领域的研究者提供了全链条的开发解决方案，有望促进该领域的算法创新和实验设计。

热度评分



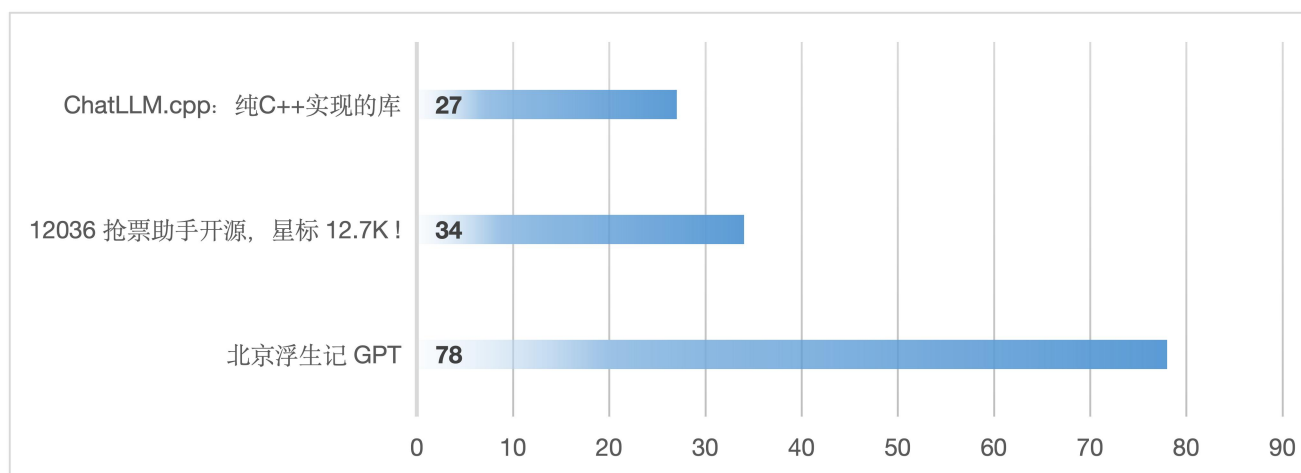
推荐关注



# 3

## 今日观察

### 1、社交观察



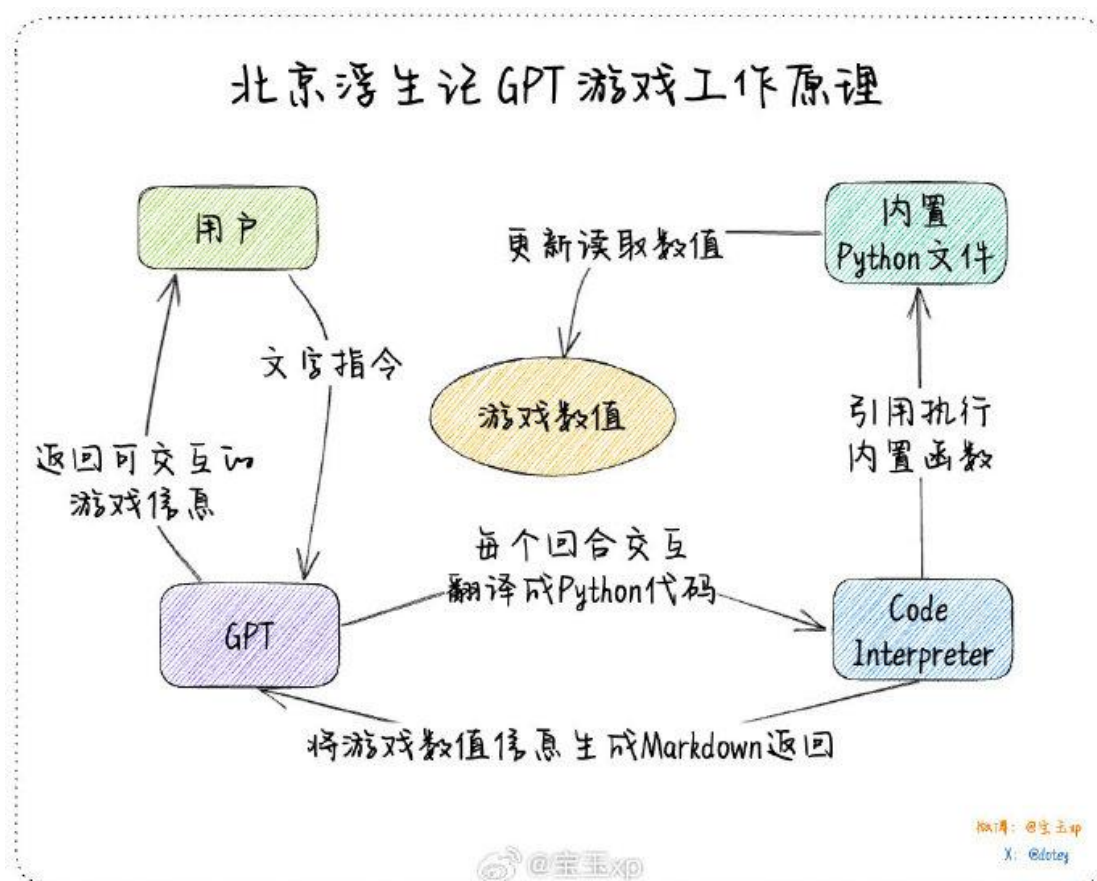
#### 1) 北京浮生记 GPT

北京浮生记 GPT 是游戏北京浮生记的 GPT 版本，让你通过 ChatGPT 也可以体验游戏。

现在正式将我的北京浮生记 GPT 开源出来了，包括了完整的 Prompt+代码+系统设计图，希望你也可以创造出属于自己的 GPT 游戏。

GPT 地址：<https://chat.openai.com/g/g-xEgLcBlnA-bei-jing-fu-sheng-ji>

项目地址：<https://github.com/JimLiu/beijing-fushengji-gpt>



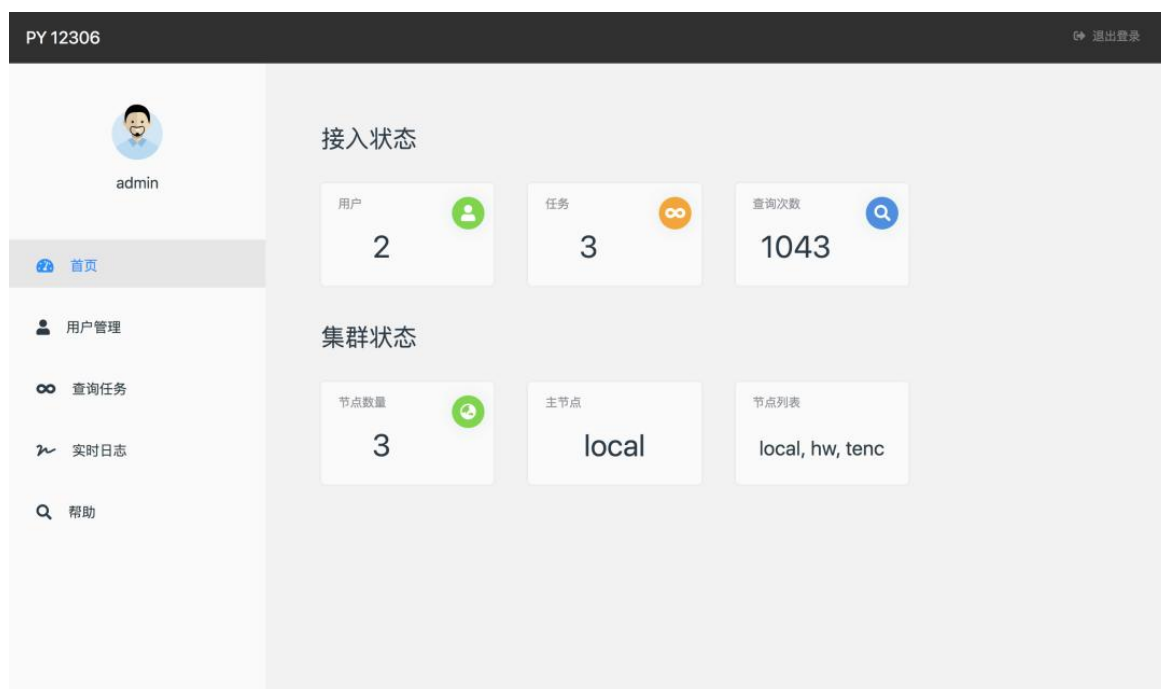
- 微博 宝玉 xp

## 2) 12036 抢票助手开源, 星标 12.7K!

这是一个 12306 购票助手, 它支持集群, 多账号, 多任务购票以及 Web 页面管理, 使用 Python 语言编写, 在 GitHub 获得了 12.7k 的 star。

开源地址: <https://github.com/pjialin/py12306>

该项目可以帮助用户在 12306 网站上更方便地购买火车票。它的特点包括支持多账号同时购票, 支持多任务购票, 以及提供一个 Web 页面来管理购票任务。



- 微信 逛逛 GitHub

### 3) ChatLLM.cpp: 纯 C++实现的库

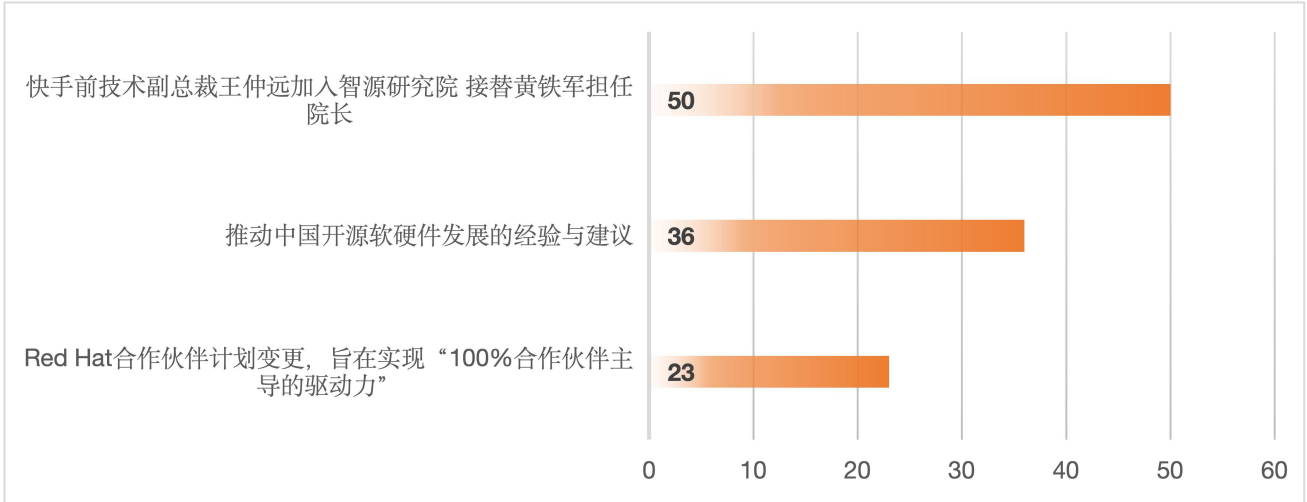
ChatLLM.cpp: 纯 C++实现的库，提供了几个模型，支持在计算机（CPU）上进行实时聊天。这些模型包括 LLaMA 系列、CodeLLaMA、Yi、WizardLM、TigerBot、Baichuan、ChatGLM、InternLM、Mistral、OpenChat、NeuralBeagle14、Phi、QWenLM、BlueLM 和 Stable-LM。这些模型经过量化，以便在 CPU 上高效运行。用户可以通过命令行界面与模型进行交互，并使用各种选项来定制聊天体验！

GitHub: [github.com/foldl/chatllm.cpp](https://github.com/foldl/chatllm.cpp)

- 微博 爱可可-爱生活



## 2、媒体观察



### 1) Red Hat 合作伙伴计划变更, 旨在实现 “100% 合作伙伴主导的驱动力”

Red Hat 即将公布的合作伙伴计划变更对于 Red Hat 来说只是一个开始, 因为它着眼于与企业 and 商业客户进行 “生态系统驱动的市场推广”, 以及围绕自动化、应用现代化以及其他业务领域实现 “100% 合作伙伴主导的驱动力和合作伙伴盈利能力” 的方法。

Red Hat 是位于美国北卡罗来纳州罗利市的 IBM 子公司和开源企业工具厂商, 该公司渠道负责人 Stefanie Chiras 在接受媒体采访时表示, 合作伙伴计划变更的其中一部分包括本季度和下季度交易注册的全球一致性。新的计划框架计划于 2024 年下半年推出, Red Hat 还为拥有先进技术知识的解决方案提供商推出了仅限邀请的合作伙伴实践加速器计划。

- 至顶网

### 2) 推动中国开源软硬件发展的经验与建议

加强开源领域的部署、投资与引导, 围绕源代码或硬件设计信息构建相适应的治理体系, 对于培育数字经济发展新动能至关重要。本文分析了开源发展的基本逻辑, 从研发支持、需求侧政策供给、开源生态建设



等方面阐述了促进开源发展的国际经验。以比较视角分析了美国开源发展现状，认为美国在开源领域具有相对较高的话语权，中国应坚持“融入全球+自主开源”双轮驱动，理清基于国外开源平台而受到制约的边界，探索开源发展战略路径，加大协同研究支持力度，开展开源部署试点并确立安全评估框架，补齐开源生态建设短板。

- 科技导报

### 3) 快手前技术副总裁王仲远加入智源研究院 接替黄铁军担任院长

2024年2月1日，北京智源人工智能研究院（下文简称：智源研究院）理事会宣布任命王仲远博士接替黄铁军，担任院长一职，全面负责研究院各项工作。据公开资料显示，王仲远现年38岁，在接棒智源研究院院长一职前，曾是快手技术副总裁，早前他还在微软亚洲研究院、Facebook、美团、脉脉等机构和科技企业工作。

王仲远在担任快手技术副总裁期间，曾参与智源大会的智能检索与挖掘论坛，并在智源社区分享了他对技术与产品应用落地的心得体会，而今接棒成为智源研究院的第二任院长。

成立于2018年11月的智源研究院，作为北京人工智能领域的新型研发机构，在国内较早系统布局起大模型技术研发、开源生态建设、人才培养和企业培育等工作。

智源研究院自成立以来，首任院长便由黄铁军担任，2023年6月，他还接任了智源研究院理事长一职。据悉，即日起，在将院长职务与王仲远完成交接后，黄铁军仍作为理事长，负责领导智源的整体战略布局和创新生态构建。

- 经济观察网

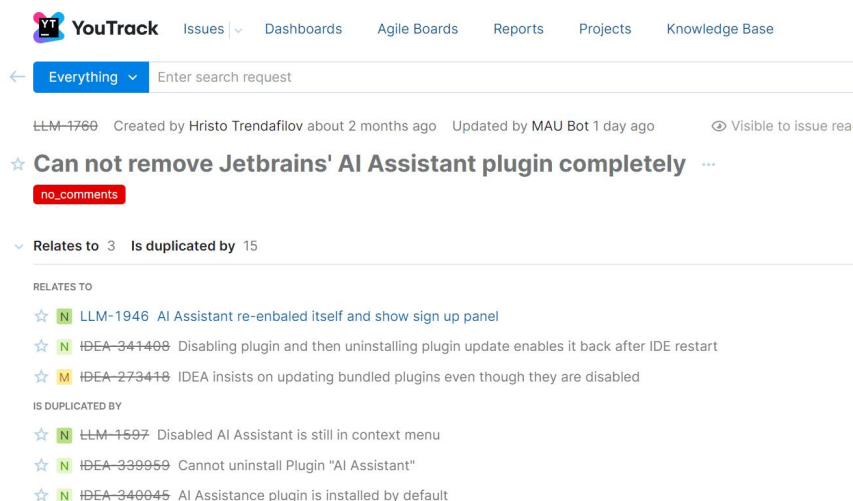
# 4

## 事件点评

### JetBrain 捆绑 AI 助手引起用户不满

JetBrains 曾于去年宣布，所有基于 IntelliJ 的 IDE 和 .NET 工具都会集成 AI 助手 (AI Assistant) 新功能；并表示会将它整合到 IDE 的核心 workflows 中，尤其是 JetBrains IDE 的强项 —— 将代码理解与 AI 深度集成。但事情并非一帆风顺，无论是出于不兼容的公司政策、还是担忧安全风险等其他原因，一些 JetBrains 用户对 AI Assistant 有着强烈的抵触情绪，他们希望在自己的 JetBrains 应用程序中不要出现该插件。然而鉴于插件代码已经 "深度集成"，相关的移除工作非常复杂。

自 AI 助手插件首次亮相以来，JetBrains 的 YouTrack 问题板上已经发布了十多个帖子，寻求删除、卸载或以其他方式删除该插件的方法。



## 事件点评

虽然部分用户对 AI 助手表示抵触，但仍有许多用户对捆绑插件感到满意。作为一个知名的开发者工具提供商，JetBrains 在开发者社区中有着广泛的影响力。他们对用户反馈持开放态度，但也需要权衡不同用户的需求。这

次决策可能对其声誉产生一定的负面影响。他们应该继续倾听用户的意见，平衡 AI 助手的功能和用户体验，以维护其在开发者社区中的良好声誉。

自身影响力	辐射影响力
★★★★☆	★★★★☆

## AI2 开源新 LLM，重新定义 open AI

艾伦人工智能研究所（Allen Institute for AI，简称 AI2）宣布推出一个名为 OLMo 7B 的新大语言模型，并开源发布了预训练数据和训练代码。OLMo 7B 被描述为 “一个真正开放的、最先进的大型语言模型”。



“有了 OLMo，open 就是真正意义上的 'open'，AI 研究界的每个人都可以访问模型创建的各个方面，包括训练代码、评估方法、数据等。AI 曾经是一个以活跃的研究社区为中心的开放领域，但随着模型的发展、成本的增加以及开始转化为商业产品，AI 工作开始闭门造车。我们希望通过 OLMo 扭转这一趋势，让研究界能够团结起来，以科学的方式更好地理解和使用语言模型，从而开发出更负责的 AI 技术，造福于每个人。”

### 事件点评

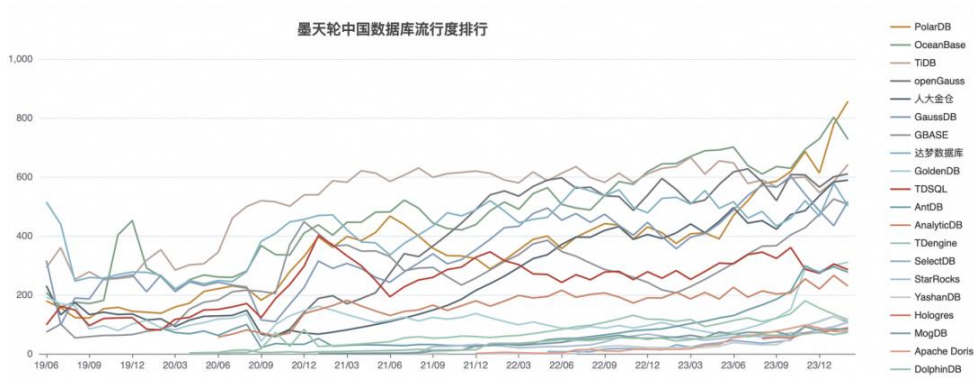
AI2 作为一个知名的人工智能研究机构，其开源的 OLMo 7B 模型将有助于推动自然语言处理领域的发展。虽然 OLMo 7B 是一个新的大型语言模型，但其开源和先进的特性将吸引广泛的研究者和开发者。AI2 的举措有

助于促进开放 AI 社区的合作和知识共享，从而推动更负责任、更先进的 AI 技术的发展。

自身影响力	辐射影响力
★★★★☆	★★★★☆

## 2024 最新一期中国数据库流行榜公布：阿里云 PolarDB 登顶

2 月 1 日，数据库社区“墨天轮”公布了 2024 年最新一期中国数据库流行度排行榜，阿里云瑶池旗下的自研云原生数据库 PolarDB 夺冠，并刷新榜单总分纪录。墨天轮评价称，因云而生的 PolarDB 数据库以客户需求为导向，并在业界开创性地落地“三层解耦”架构、多主多写、HTAP、Serverless、集中与分布式一体化等一系列业内领先的前沿技术，为用户提供了灵活、可靠的数据库服务，解决了众多业务问题。



### 事件点评

这一事件表明 PolarDB 在中国数据库领域的影响力和认可度持续增强。不过虽然 PolarDB 在中国数据库领域的影响力已经很大，但仍有一些其他数据库产品如 OceanBase、TiDB、OpenGauss 等也在榜单中有不俗的表现。因此，尽管 PolarDB 夺冠，但在整个数据库市场中的辐射影响力还有提升的空间。

自身影响力	辐射影响力
★★★★☆	★★★★☆

# 5

今日推荐

## myshell-ai/OpenVoice

### 项目介绍:

OpenVoice, 这是一种多功能的即时语音克隆方法, 只需要参考说话者的一个简短的音频剪辑即可复制他们的声音并生成多种语言的语音。

除了复制参考说话者的音色之外, OpenVoice 还可以对语音风格进行精细控制, 包括情感、口音、节奏、停顿和语调。OpenVoice 还可以针对海量说话人训练集中未包含的语言实现零样本跨语言语音克隆。OpenVoice 的计算效率也很高, 其成本比性能较差的商用 API 低数十倍。



Stars	Forks	Watching
13.5k	1.2k	126

### 推荐理由:

OpenVoice 是一款强大且灵活的语音克隆工具, 有着精准的语言克隆、灵活的语音风格控制、同时满足高效且低成本; 适用于各种语音应用场景。

# 6

## 每日项目榜

### GitHub 榜单

Rank	Language	Stars	Forks	Stars Today
01	danielmiessler/fabric			
	fabric is an open-source framework for augmenting humans using AI.			
	Python	2568	211	1139
02	InkboxSoftware/excelCPU			
	16-bit CPU for Excel, and related files			
	Python	2523	181	337
03	f/awesome-chatgpt-prompts			
	This repo includes ChatGPT prompt curation to use ChatGPT better.			
	HTML	98947	13607	115
04	all-in-aigc/aicover			
	ai cover generator			
	TypeScript	979	166	205
05	facebookresearch/codellama			
	Inference code for CodeLlama models			

	Python	12981	1330	197
06	webprodigies/plura-production			
	TypeScript	317	149	43
07	ExOK/Celeste64			
	A game made by the Celeste developers in a week(ish, closer to 2)			
	C#	901	62	170
08	haotian-liu/LLaVA			
	[NeurIPS'23 Oral] Visual Instruction Tuning (LLaVA) built towards GPT-4V level capabilities and beyond.			
	Python	13178	1372	155
09	mlflow/mlflow			
	Open source platform for the machine learning lifecycle			
	Python	16585	3814	127
10	PKU-YuanGroup/MoE-LLaVA			
	Mixture-of-Experts for Large Vision-Language Models			
	Python	697	35	219

# 7

## 用户观点

[思考]备选#人工智能[超话]

有趣

我司在开源的基础上，半年下来做了大量的工作才超过GPT4，因为用了开源技术这个性质，故对外做一些开放，我觉得也是应该的终于来了

感觉有webagent以后这些都不太需要了？

我也做了一个

数据开源

### 1、Visual Studio Code 1.86 发布

- 观点 1：这个版本升级了最小 glibc 版本，必须 $\geq 2.28$ ，所以用 ubuntu18.04 的需要注意了。
- 观点 2：隔壁大 JetBrains 加了个 AI 助手，都炒翻天了，这边就没一点波澜？

### 2、MinIO 权限提升漏洞

- 观点 1：啥都高性能，文章天天瞎写
- 观点 2：你还别说，minio 性能还真高！我们项目里就集成了，效果很好。

### 3、还在争论 jpa 和 mybatis? 国产 Java ORM: sqltoy-orm 5.2.92 发版

- 观点 1：“通常的项目用 jpa 即可，个别项目可以用 mybatis!
- 观点 2：2024 要换思维了，要兼顾 2 者优势解决项目中各种问题

### 4、禅道 20.0.beta1 发布啦，重构底层 PHP 和 UI 框架，用户体验全面升级

- 观点 1：禅道对于产品开发还是很实用的。



# 8

## 媒体观点

### 1、钛媒体

**观点：数字员工、超级个体、具身智能，AI Agent 未来发展十大研究方向**

埃森哲在《技术展望 2024》报告中指出，96%的企业高管认为 AI Agent 生态系统应用将在未来 3 年内为他们的组织带来重大机遇。报告认为，随着人工智能向智能体演进，自动化系统将能够自主决策和行动。智能体不仅会为人类提供建议，还将代表人类采取行动。人工智能将继续生成文本、图像和洞察，而 AI Agent 将自行决定如何处理这些信息。

当智能体升级成人类的同事后，就需要人类与智能体一起重新构建技术和人才的未来。

### 2、财经十一人

**观点：谷歌营收增速乏力，没有吃到大模型第一波红利**

谷歌的想象空间在缩小。相比形成了“云-软件-AI”三条轮动增长曲线的微软，谷歌主营业务融合不佳。相比垄断手机市场 80%以上利润的苹果，谷歌的广告业务统治力不够强，正面临一批新的挑战者。谷歌三大业务分别是广告、订阅平台和设备、云。谷歌增速探底，核心原因是，广告、云两大业务的营收增速都在不断下滑。

### 3、凤凰网科技

**观点：大模型训练可“自动驾驶”，蚂蚁集团全面开源相关技术**

近日，蚂蚁集团 AI 创新研发部门 NextEvo 全面开源 AI Infra 技术，可帮助大模型千卡训练有效时间占比超过 95%，能实现训练时“自动驾驶”，这推动了 AI 研发效率。该技术框架名为 DLRouter，目标在于大规模分布式训练的智能化。目前很多企业的训练作业都是跑在混合部署的集群中，运行环境复杂多变，不管多么“崎岖的地形”，DLRouter 都可以“轻松行驶”。

## 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。