

# 开源日报

OPENSOURCE DAILY



Issue: 26.01.2024

Date: 20.02.2024



总第012期

开源中国社区 - 数据分析



# ABOUT US

## 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。

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。

**OSChina 数据分析部**

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

<b>开源数据   OPEN SOURCE DATA</b>	
开源社会数据	
开源环境指数	
开源话题情绪走势	
开源话题热词走势	
渠道传播TOP	
今日活跃度渠道评估	
开源话题区域提及走势	
<b>今日要点   TODAY FOCUS</b>	
Linux 内核成为 CVE 编号机构 (CNA)	
Dart 3.3 发布: 扩展类型、JavaScript Interop	
<b>今日观察</b>	
社交观察	
媒体观察	
<b>事件点评</b>	
2026 年传统搜索引擎流量将下降 25%	
字节跳动推出中文版 “Sora” ? 回应来了	
<b>今日热门项目</b>	
Plachtaa/VALL-E-X	
<b>每日Github精选</b>	
<b>每日Gitee精选</b>	
<b>开源之声</b>	
用户观点	
媒体观点	
<b>附: 中国基础软件开源产业主要参与者</b>	
<b>附: 中国基础软件开源产业主要参与者</b>	
<b>报告声明</b>	

# 开源数据 | OPEN SOURCE DATA

## 开源社会数据

2024年02月19日 18:00:00~2024年02月20日 18:00:00, 共监测到信息总量5857条, 其中网页1359条 (占比23.2%), APP1199条 (占比20.47%), 微信1061条 (占比18.11%), 微博644条 (占比11%), 视频571条 (占比9.75%), 论坛398条 (占比6.8%), 问答332条 (占比5.67%), 头条号289条 (占比4.93%), 报刊4条 (占比0.07%)。



### 新闻平台

数据: 1359条 占比: 23.2%



### 社交平台

数据: 1705条 占比: 29.11%



### 视频平台

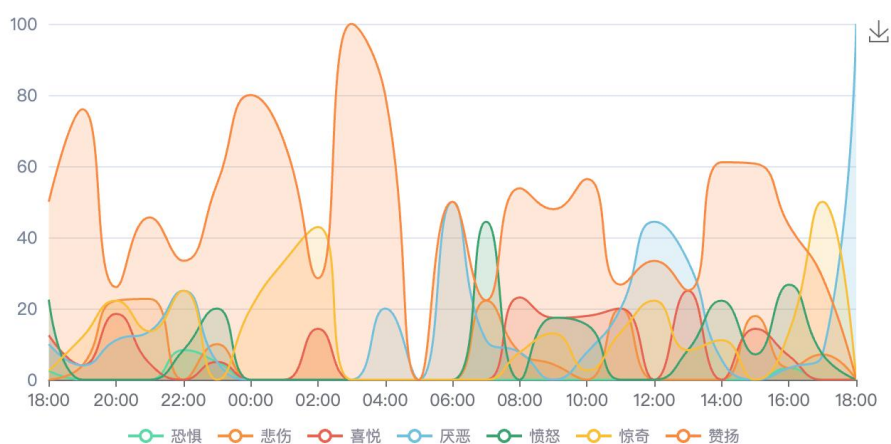
数据: 571条 占比: 9.75%



### APP平台

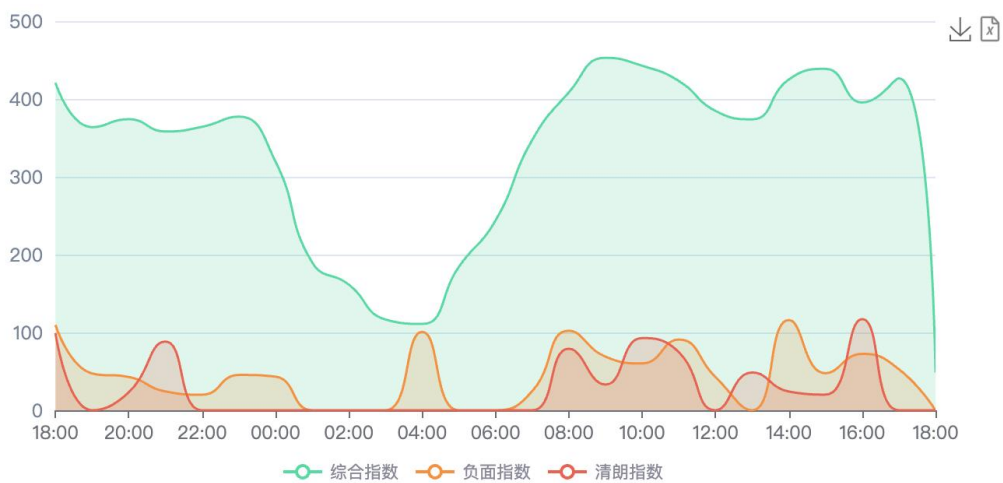
数据: 1199条 占比: 20.47%

## 开源话题情绪走势



今日开源环境舆论中，喜悦和赞扬占比6成左右，情绪一般

## 开源环境指数



舆情指数模型POI，根据舆情系统目前可采集数据设定。针对正面舆情跨平台传播热度进行评估（网页、微信、微博、客户端文章数量和正面舆情数量占比），暂时未考虑媒体影响力、事件敏感程度、历史风险性等其他要素。

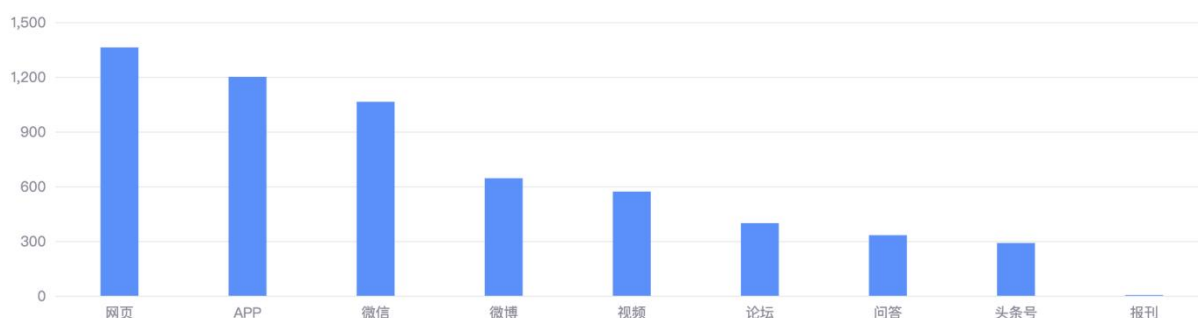
## 开源话题热词走势



分析时段内，“模型”一词出现频率最高，达1815次，其次是“开源”，出现次数达1769次。

今日开源话题热词中，“LLM”“模型”“Sora”等关键词在社交中热度较高

## 渠道传播TOP



分析时段内，网页信息1361条(占比23.22%)、APP信息1200条(占比20.47%)、微信信息1063条(占比18.13%)、微博信息644条(占比10.99%)、视频信息571条(占比9.74%)、论坛信息398条(占比6.79%)、问答信息332条(占比5.66%)、头条号信息289条(占比4.93%)、报刊信息4条(占比0.07%)。

地区	数量
北京	770
广东	350
上海	248
河南	58
江苏	39
浙江	35
湖北	33
四川	29
河北	27
安徽	27

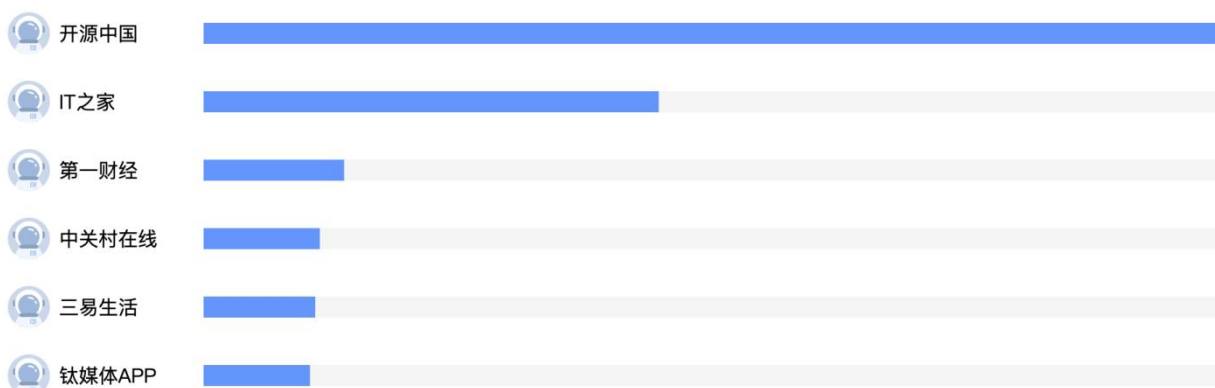


## 开源话题区域提及走势

分析时段内，发布信息地域信息地域主要集中在北京770条（占比43.58%）、广东350条（占比19.81%）、上海248条（占比14.04%）、河南58条（占比3.28%）、江苏39条（占比2.21%）。

今日北京、广东、上海地区对开源话题的兴趣度占比在 77% 左右。

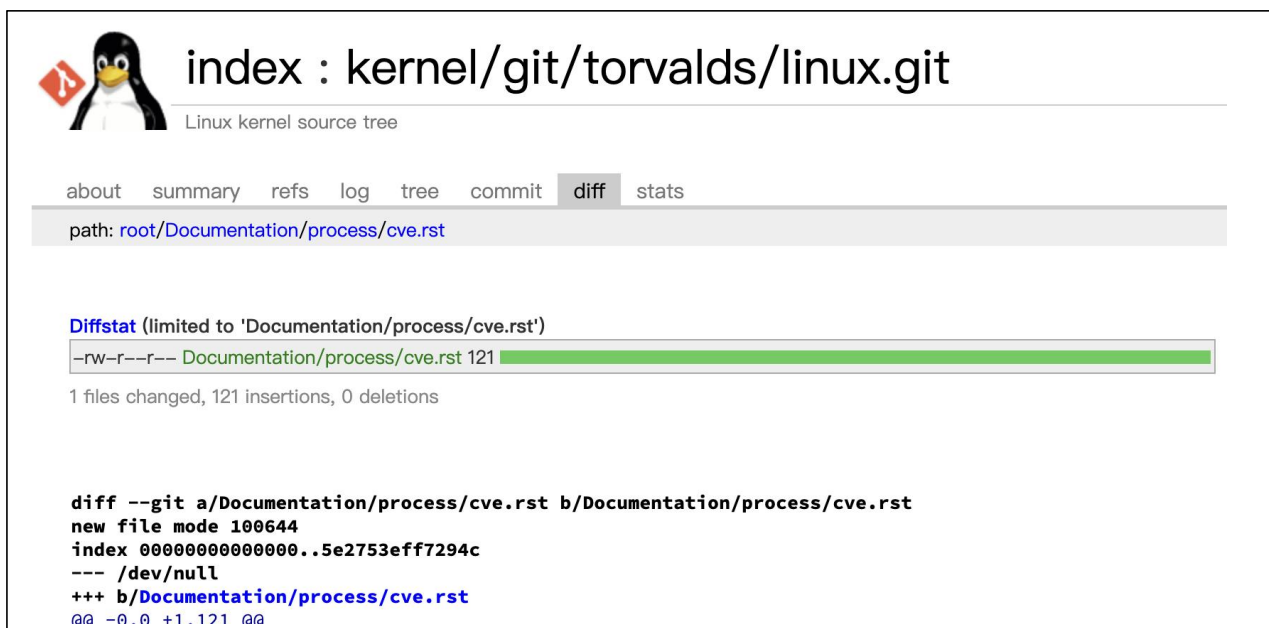
## 今日活跃度渠道评估



## Linux 内核成为 CVE 编号机构 (CNA)

Linux 内核已被接受为 CVE 编号机构 (CNA)，这意味着他们将直接管理内核的 CVE。Linus Torvalds 近日在邮件列表发布了 Linux 6.8-rc5，并介绍称文档添加了 CVE 漏洞处理相关的指南。

文档写道，Linux 内核开发团队有能力为潜在的内核安全问题分配 CVE，而分配的 CVE 编号将在 linux-cve-announce 邮件列表上公布，修复的安全漏洞才会分配 CVE 编号，未修复的不会自动分配编号。



```
index : kernel/git/torvalds/linux.git
Linux kernel source tree

about summary refs log tree commit diff stats
path: root/Documentation/process/cve.rst

Diffstat (limited to 'Documentation/process/cve.rst')
-rw-r--r-- Documentation/process/cve.rst 121

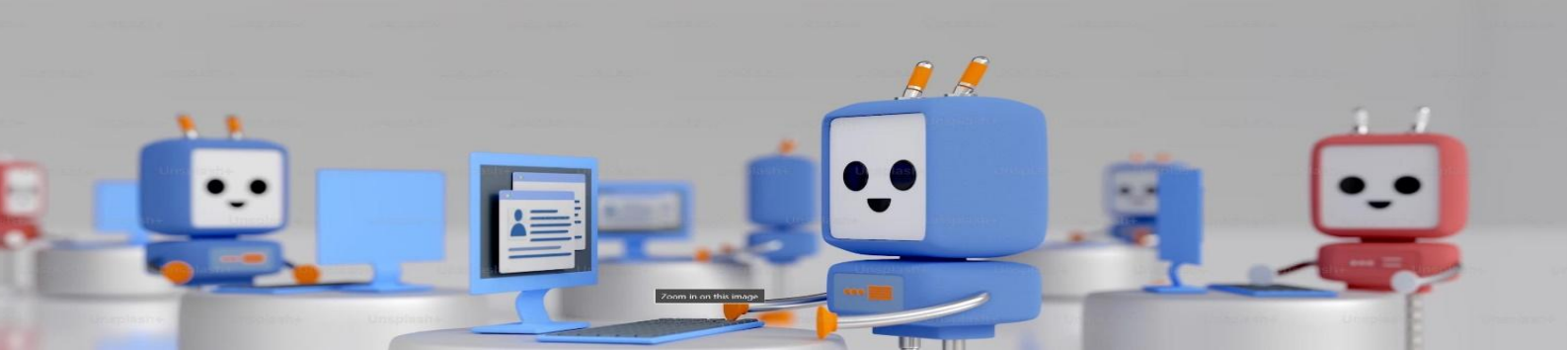
1 files changed, 121 insertions, 0 deletions

diff --git a/Documentation/process/cve.rst b/Documentation/process/cve.rst
new file mode 100644
index 00000000000000000000000000000000..5e2753eff7294c
--- /dev/null
+++ b/Documentation/process/cve.rst
@@ -0,0 +1,121 @@
```

### 编辑简评

Linux 内核作为开源项目的重要代表之一，其被接受为 CVE 编号机构无疑是一个里程碑式的事件。这一举措意味着 Linux 内核团队将直接管理内核相关的 CVE，有助于更及时地发现、报告和修复潜在的安全漏洞，进一步提升了 Linux 内核的安全性和可信度。

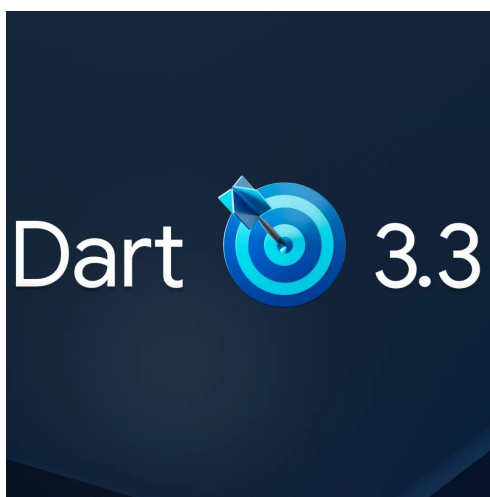




# Dart 3.3 发布：扩展类型、JavaScript Interop

Dart 3.3 现已发布，公告称此版本改变了性能和跨平台开发的游戏规则。

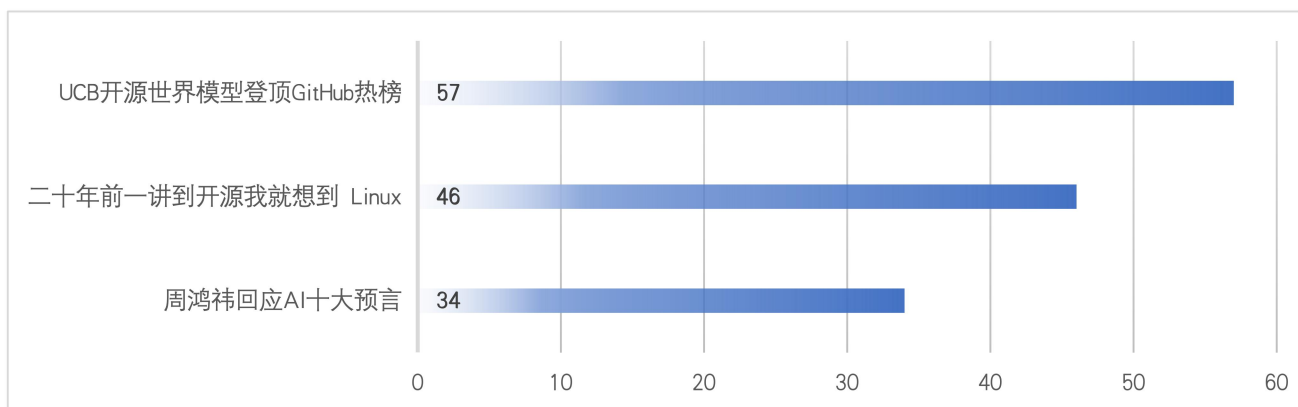
增强的扩展类型（Extension Types）将彻底改变性能优化以及用户与本地代码的交互方式。JavaScript interop 模型也得到了改进，引入了强大的类型安全性和开发人员友好的方式来利用 Web 平台的强大功能。“所有这些都为 WebAssembly 支持铺平了道路”。此外，新版本还增加了 Google AI 功能。



## 编辑简评

新版本引起了相当的关注，尤其是对于那些关注性能和跨平台开发的开发者来说。增强的扩展类型和改进的 JavaScript interop 模型为 Dart 的性能和开发体验带来了显著提升，而新增的 Google AI 功能也为开发者提供了更多的工具和资源。整体而言，这次更新为 Dart 的未来发展打下了坚实的基础，对于推动跨平台应用和 WebAssembly 的发展具有重要意义。

### 社交观察



#### 周鸿祎回应AI十大预言

周鸿祎在被问到现在怎么看年初提出的AI十大预言时，回应称已经实现4个了。分别是开源大模型爆发、大模型运行在终端、文生图和文生视频取得突破性进展、多模态成为大模型标配。周鸿祎认为，特别是Sora的出现让他感到很意外，发展得很快。

- 微博 中国企业家杂志

#### 二十年前一讲到开源我就想到 Linux

二十年前一讲到开源我就想到 Linux，现在讲到开源我就想到如果一个软件不开源估计就没什么人用了。也许这就是二十年来软件业的进步吧？

- 微博 2gua

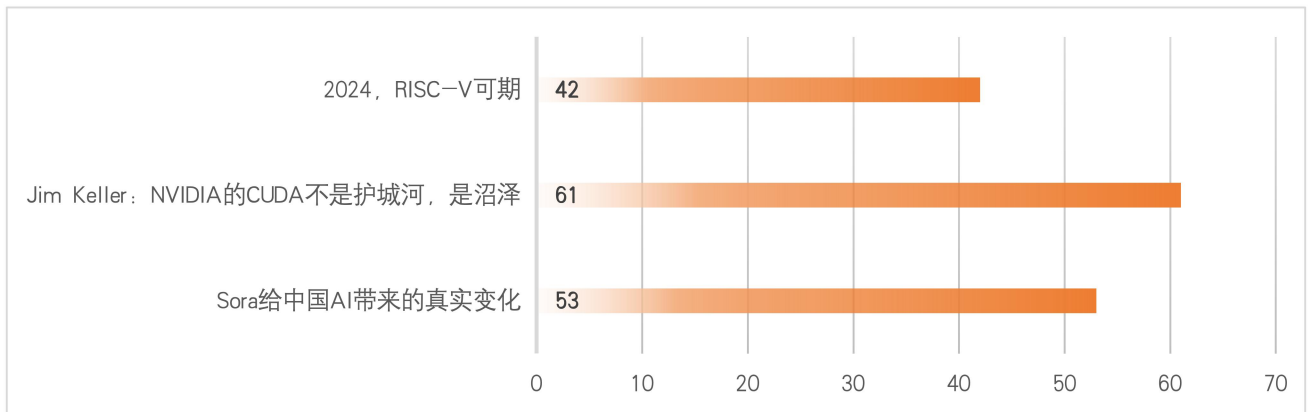
#### UCB开源世界模型登顶GitHub热榜

今日GitHub热榜榜首，是最新的开源世界模型。上下文窗口长度达到了100万token，持平了谷歌同时推出的王炸Gemini 1.5，伯克利出品。

强大的模型，命名也是简单粗暴——没有任何额外点缀，直接就叫LargeWorldModel（LWM）。

- 微博 量子位

媒体观察



### Sora给中国AI带来的真实变化

国产 Sora 一定会出现，但我们做好各行业规模应用 Sora 的准备了吗？恐怕今天的答案还是，没有。

前面提到，Sora 的 "暴力美学" 再次证明了 Scale 的价值。而要达到涌现效果，基座模型仍然高度依赖于大量高质量数据集，超大规模算力，大量工程化调优人才，以及由此带来的巨大开发及运行成本。

- 脑极体

### Jim Keller: NVIDIA的CUDA不是护城河，是沼泽

对于 x86、Arm、MISC 和 RISC-V 等处理器架构都有深入研究的传奇处理器架构师Jim Keller在本周末批评了被外界认为是英伟达（NVIDIA）“护城河”的 CUDA架构和软件堆栈，并将其比作x86，称之为“沼泽”。他指出，就连英伟达本身也有多个专用软件包，出于性能原因，这些软件包依赖于开源框架。“CUDA 是沼泽，而不是护城河，”

- 芯智讯

### 2024，RISC-V可期

2023年，长期被冠以“低端”帽子的RISC-V架构，终于实现了高端化过程中的两个“小目标”：一个是单核性能走高，可与ARM Cortex-A7对标；另一个是应用场景拓展到PC领域，首台搭载RISC-V架构的笔记本电脑面世。

如此势头下，2024年RISC-V的发展似乎“一片坦途”。

- 中国电子报

# 事件点评

## 2026 年传统搜索引擎流量 将下降 25%

### Gartner Predicts Search Engine Volume Will Drop 25% by 2026, Due to AI Chatbots and Other Virtual Agents

Tech Marketing Tactics to be Discussed at the Gartner Tech Growth & Innovation Conference, March 20-21 in Grapevine, Texas

By 2026, traditional search engine volume will drop 25%, with search marketing losing market share to AI chatbots and other virtual agents, according to Gartner, Inc.

"Organic and paid search are vital channels for tech marketers seeking to reach awareness and demand generation goals," said Alan Antin, Vice President Analyst at Gartner. "Generative AI (GenAI) solutions are becoming substitute answer engines, replacing user queries that previously may have been executed in traditional search engines. This will force companies to rethink their marketing channels strategy as GenAI becomes more embedded across all aspects of the enterprise."

#### Quality and Authenticity Will Be Focal Points as Virtual Agents Replace Traditional Search

With GenAI driving down the cost of producing content, there is an impact around activities including keyword strategy and website domain authority scoring. Search engine algorithms will further value the quality of content to offset the sheer amount of AI-generated content, as

市场分析公司 Gartner 指出，生成式 AI 已严重威胁传统搜索引擎。预测到 2026 年，传统搜索引擎的流量将下降 25%，搜索营销的市场份额将被人工智能聊天机器人和其他虚拟代理所取代。

Gartner 副总裁分析师 Alan Antin 称，有机搜索和付费搜索是科技营销人员实现认知和需求生成目标的重要渠道。GenAI 解决方案正在成为替代答案引擎，取代以前可能在传统搜索引擎中执行的用户查询。随着 GenAI 越来越深入企业的方方面面，这将迫使企业重新思考其营销渠道战略。公司需要专注于制作对客户和潜在客户有用的独特内容。



### 事件点评 Event Search

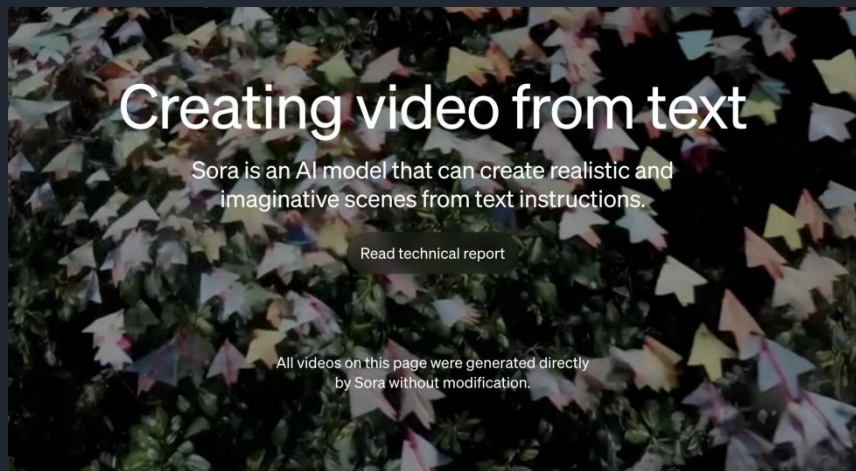
搜索引擎在人们的日常生活中扮演着重要角色，因此其变化会引起广泛关注。预计搜索量下降将影响广告、营销和用户行为，对整个科技产业产生影响。

自身影响力



辐射影响力





## 字节跳动推出中文版“Sora”？回应来了

近日，市场有消息称，在 Sora 引爆文生视频赛道之前，国内的字节跳动也推出了一款新型视频模型 Boximator，与 Gen-2、Pink1.0 等既有模型相比，Boximator 的独特之处在于，它能够通过文本输入精确地控制视频中人物或物体的运动。

2 月 20 日，字节跳动相关人士回应称，Boximator 是视频生成领域控制对象运动的技术方法研究项目，目前还无法作为完善的产品落地，距离国外领先的视频生成模型在画面质量、保真率、视频时长等方面还有很大差距。

Sora 可根据文本描述生成长达 60 秒的视频，其中包含精细复杂的场景、生动的角色表情以及复杂的镜头运动。



### 事件点评 Event Search

字节跳动作为一家知名科技公司，其推出的新产品自然会引起关注。然而，由于Boximator尚未成熟，其影响力有限；字节跳动需要进一步优化它，以满足用户的需求。

自身影响力

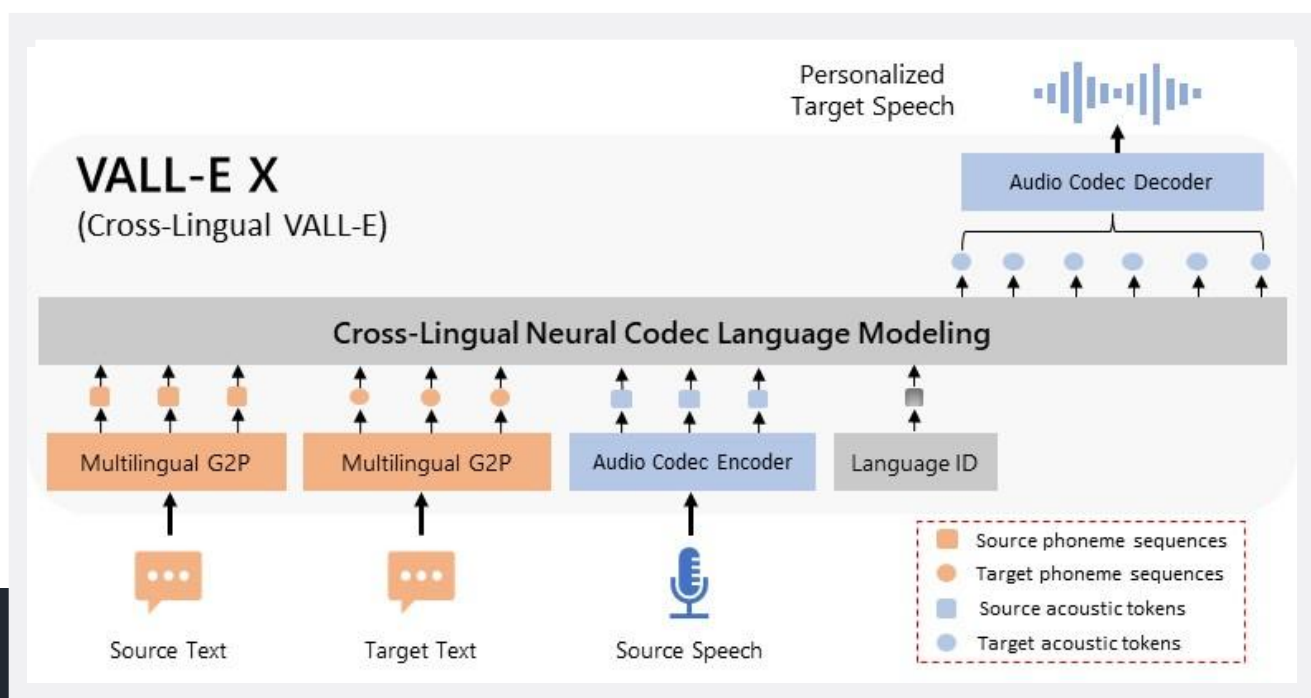


辐射影响力



# 今日热门项目

## Plachtaa/VALL-E-X



### 项目介绍

VALL-E X 是一个强大而创新的多语言文本转语音 (TTS) 模型，最初由微软发布。VALL-E X 可以为单语使用者合成另

一种语言的个性化语音。以源自源文本和目标文本的音素序列以及源自音频编解码器模型的源声学标记作为提示，

VALL-E X 能够生成目标语言的声学标记，然后将其解压缩为目标语言语音波形。

### 推荐理由

VALL-E X 是一款强大且创新的多语言文本到语音合成与语音克隆模型。它支持中文、英文和日语，具有零样本语音克隆

功能，能够生成高质量的个性化语音。此外，它还具备口音控制和声学环境维护功能，为开发者提供了灵活且自然的语

音合成体验。

Stars

6.9k

Forks

670

Watching

77

# 每日Github精选

每日 Github 精选将为您展示来自全球开发者的创意与创新，我们精选的项目不仅具有实用价值，更体现了开源社区的活力与智慧。无论是寻找灵感、学习新技术，还是了解行业动态，这个栏目都能为您提供丰富的资源。

Rank	Language	Stars	Forks	Stars Today
01	karpathy/minibpe Minimal, clean, code for the Byte Pair Encoding (BPE) algorithm commonly used in LLM tokenization.			
	Python	4687	295	1372
02	LargeWorldModel/LWM			
	Python	4518	341	1882
03	RVC-Boss/GPT-SoVITS 1 min voice data can also be used to train a good TTS model! (few shot voice cloning)			
	Python	14195	1496	828
04	facebookresearch/DiT Official PyTorch Implementation of "Scalable Diffusion Models with Transformers"			
	Python	2842	233	386
05	ChatGPTNextWeb/ChatGPT-Next-Web A cross-platform ChatGPT/Gemini UI (Web / PWA / Linux / Win / MacOS).			
	TypeScript	61084	52395	125
06	matter-labs/zksync-era zkSync era			
	Rust	1931	1568	104
07	krahets/hello-algo 《Hello 算法》：动画图解、一键运行的数据结构与算法教程，支持 Python, C++, Java, C#, Go, Swift, JS, TS, Dart, Rust, C, Zig 等语言。English edition ongoing			
	Java	62777	7760	159
08	google/magika Detect file content types with deep learning			
	Python	5150	243	2682
09	AUTOMATIC1111/stable-diffusion-webui Stable Diffusion web UI			
	Python	121756	24250	293
10	jlevy/the-art-of-command-line Master the command line, in one page			
		144930	14225	467

# 每日Gitee精选

每日Gitee精选将为您展示来自国内开发者的创意与创新，我们精选的项目不仅具有实用价值，更体现了开源社区的活力与智慧。无论是寻找灵感、学习新技术，还是了解行业动态，这个栏目都能为您提供丰富的资源。

Rank	Language	Stars	Forks
01	MarsTech/Mars-Admin 管理后台脚手架： Spring Cloud微服务 + Spring Security SSO + JWT Token +自定义登陆页面 + 前后端分离 + Ant Design Pro + K8s部署 +Swagger + 多数据源 + Redis		
	Java	4.5k	441
02	美乐低代码平台 / MetaLowCode 企业级全栈低代码平台 企业级全栈低代码开发平台！ 潜心打造强大核心技术底座： 内置元数据引擎、业务流程引擎、业务触发器、可视化表单、仪表盘设计、智能报表引擎、数据挖掘和插件扩展框架		
	JavaScript	2k	505
03	LvHuaiSheng / gzm-design 一款漂亮易用的在线设计器，支持PSD导入、PSD解析，可用于海报设计器、广告设计器、logo设计器、AI创作图片合成器等		
	TypeScript	60	19
04	Vanishi / BXC_VideoAnalyzer_v4 基于C++开发的视频行为分析系统v4系统，可以在不用考虑音视频开发，编解码开发，界面开发等情况下，只需要训练自己的模型，开发自己的算法插件，就可以轻松实现出任何想要的视频行为检测，比如周界入侵，烟火检测，打架，斗殴，跌倒，人群聚集，电动车，垃圾箱，抽烟，攀爬，离岗睡岗，安全帽，充电桩，工作服，疲劳检测，交通拥堵等等。		
	C++	273	137
05	网易有道 / QAnything 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		
	Python	94	11
06	WuKongDev / WuKongIM 高性能通用通讯服务，支持即时通讯（聊天软件）(IM)(Chat)，消息推送，消息中台，物联网通讯，音视频信令，直播弹幕，客服系统，AI通讯，即时社区等场景		
	Go	324	94
07	dotNET China / Furion 让 .NET 开发更简单，更通用，更流行。		
	C#	10.8k	4k
08	huhai / YHH YHH比赛项目代码存放仓库		
	Vue	10	4



# 开源之声

在开源的世界里，每一个声音都值得被聆听。为了更好地传递和汇聚开源社区的多元观点，我们特别推出了“每日开源之声”栏目。这个栏目的宗旨是捕捉和分享开源社区内关于最新动态、内容、产品、新闻等话题的用户发声。

“每日开源之声”将每日精选来自国内各大社交平台的热点讨论、用户评论和专业分析，为您提供第一手的开源情报和深度的观点碰撞。我们关注的不仅是技术本身，更关注技术背后的思想火花和社会影响。

## 用户观点

Windows 10停更导致2.4亿台PC淘汰，谷歌：快用ChromeOS Flex “重生”

观点1：如果有选择，我绝对放弃Windows

观点2：○【fuchsia】邀请您参与语音通话

× 拒绝    ✓ 下次再接听

“我第一次用 GitHub，别 TM 给我代码，我只要安装包”

观点1：问题在于 Github 本来不就是面向开发者的么？;-)

观点2：第一次钓鱼，别TM给我鱼竿，我只要鱼。

观点3：所以将开源软件打包到应用商店，是真的有市场。

TIOBE 2 月：Go 首次进入前十、“上古语言” COBOL 和 Fortran 排名飙升

观点1：Fortran 挺方便的

观点2：Java 最流行的 ORM 竟然是基于 XML 的 Mybatis，都 2024 年了，2024 年了啊，还在

XML 里面拼 SQL，简直不可思议！！

# 媒体观点

## 钛媒体APP

### 中国大模型赛道最大单笔融资，月之暗面获 10 亿美金投资，投后估值约 25 亿美金

近日，大模型创业公司月之暗面完成新一轮超10亿美金的融资，投资方包括红杉中国、小红书、美团、阿里等，上一轮的老股东继续跟投，本轮融资后月之暗面估值已达约25亿美金。

这是自大模型创业潮兴起以来，中国大模型赛道金额最大的一笔单轮融资。

## 中国企业家杂志

### 周鸿祎回应 2024 年 AI 行业十大预言，现已实现四个

中国企业家杂志独家对话了知名企业家、360 集团董事长周鸿祎，共同探讨了近期备受瞩目的 AI 技术 Sora 所带来的重大影响。在被问及如何看待年初提出的 AI 十大预言时，周鸿祎兴奋地表示，目前已经有四个预言得以实现，分别是开源大模型爆发、大模型运行在终端、文生图和文生视频取得突破性进展、多模态成为大模型标配。周鸿祎认为，特别是 Sora 的出现让他感到很意外，发展得很快。

## 环球网

### 华为发布全新数据湖解决方案及全闪存新品

在数字经济的快速发展中，数据已成为关键的生产要素。释放数据的价值对于经济的增长至关重要。为此，2月20日，华为推出了新的数据湖解决方案和全闪存新品，旨在助力金融、政府、运营商及教育科研等行业充分挖掘数据资产的潜能，推动数据要素市场的构建和数据资产化进程。

## 附：中国基础软件开源产业主要参与者

### 基础软件开源项目

#### AI框架



#### 大模型



#### 中间件



## 数据库



## 操作系统



## 开源组织

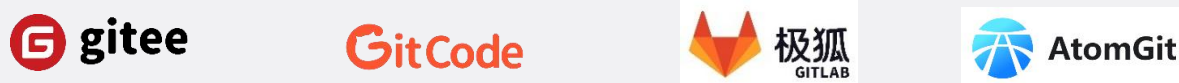
### 开源社区评估机构



### 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。

# 关键数据来源

所有数据来源于公开数据，包括但不止于媒体门户、技术社区、问答平台、技术论坛、微博、博客、知乎、Gitee、Github、抖音、小红书、B 站、微信公众号、报刊等渠道。