

第18期--无界雕塑作品

一、刊首图



甘肃省瓜州县的戈壁滩上，有一座6000多根白色钢管搭建的雕塑作品[《无界》](#)，象征唐代的边关

二、时事新闻

1、[紫金山-阿特拉斯彗星](#)

紫金山-阿特拉斯彗星（国际编号C/2023 A3）10月9日进入近地点，直到26日离开。这段时间内，肉眼可以清晰看到。这几天，世界各地的天文爱好者拍下了许多照片。它的轨道周期是6万多年，上次光临地球的时候，人类还是石器时代。



2、[ChatGPT 4o with canvas](#)

2024.10.14, OpenAI 推出了一项名为“Canvas”的新功能，它提供了一种新的交互方式，用于与ChatGPT在写作和编程项目上进行更好的协作，比如选定内容，进行针对性修改。

The screenshot shows the ChatGPT 4o with canvas interface. On the left, there is a chat history with a message from the user asking for a project plan and a response from ChatGPT providing a framework. The main area displays a document titled "大模型项目建设方案" (Large Model Project Construction Plan) with the following sections:

- 一、项目背景** (Project Background): Discusses the rapid development of AI and the need for a project based on large models.
- 二、项目目标** (Project Objectives):
 - 提升智能服务能力: 利用大模型进行自然语言处理, 实现高效、精准的自动化问答和文本生成。
 - 优化工作流程: 通过大模型技术辅助行业专家, 减轻重复性劳动, 提升工作效率。
 - 推动技术创新: 以大模型为核心技术, 探索行业内新的业务模式和应用场景。
- 三、项目范围** (Project Scope):
 - 数据准备: 采集并整理训练所需的海量高质量数据, 包括结构化数据和非结构化文本数据。
 - 模型训练与微调: 基于现有大模型进行微调, 确保模型适应特定领域需求。
 - 系统集成: 将大模型与现有系统 (如客服系统、文档管理系统等) 进行集成, 确保数据和功能的无缝连接。
 - 应用开发: 开发用户友好的应用接口, 使用户能够便捷地使用大模型提供的智能服务。
- 四、实施计划** (Implementation Plan):
 - 阶段一: 需求分析与可行性评估 (3个月)
 - 调研用户需求, 明确大模型应用场景。
 - 对现有的硬件和软件环境进行评估, 制定可行性报告。
 - 阶段二: 数据准备与模型选型 (4个月)
 - 收集与整理行业相关的数据, 进行数据清洗和标注。
 - 确定合适的大模型架构, 并配置训练环境。

3、[诺贝尔物理学奖颁发给人工智能领域科学家](#)

2024.10.8, 2024年诺贝尔物理学奖揭晓。John J. Hopfield、Geoffrey E. Hinton获奖，以表彰他们在“基于人工神经网络的机器学习的基础性发现和发明”做出了重大贡献。

为什么将诺贝尔物理学奖授予研究人工神经网络的学者？诺贝尔奖委员会解释，是因为人工神经网络是用物理学工具训练的。也就是说，目前红遍全世界的人工智能的原始基础是物理学的发现和知识，机器学习模型是基于物理原理实现的。

4、[AlphaFold2预测蛋白质结构获得诺贝尔化学奖](#)

2024.10.9, 2024年诺贝尔化学奖授予了三位科学家, 表彰他们在蛋白质设计和预测领域的突破性贡献。美国科学家 David Baker 因成功构建全新蛋白质获得一半奖金; 英国科学家 Demis Hassabis 和美国科学家 John Jumper 因开发 AI 模型 AlphaFold2, 解决蛋白质结构预测的难题, 共享另一半奖金。此次奖项强调了蛋白质作为生命重要化学工具的关键性作用, 得主们破解了蛋白质复杂结构的密码, 推动了这一领域的发展。

5、[OpenAI发布ChatGPT网络搜索](#)

2024.11.1, OpenAI正式发布了ChatGPT网络搜索功能, 能实时、快速获取附带相关网页来源链接的答案。



6、[谷歌Willow量子芯片](#)

2024.12.10, 谷歌最新一代量子芯片Willow炸裂推出, 最大突破在于超强的计算能力和纠错能力。对于一项名为“随机电路采样”的基准任务, 当前最快的超级计算机需要花费10的25次方年来解决, 时长远超宇宙年龄(267亿年), 而Willow完成这项任务的时间不到5分钟。

7、[树莓派掌上电脑](#)

一支国外团队推出了“树莓派掌上电脑”, 它就是一个全功能的 Linux 系统, 带有7小时续航。它还有一个可组装版本, 屏幕是单独分离的, 下面可以配上键盘, 也可以配上手柄。它是开源的, 制作团队希望将成本控制在250美元。



8、[GitHub Copilot免费向所有用户开放](#)

2024.12.19, GitHub宣布, 其 AI 编程助手 GitHub Copilot 推出免费计划, 现已在 Visual Studio Code 中向所有用户开放。用户只需拥有 GitHub 账户即可开始使用, 无需信用卡或订阅流程。然而, 免费版本存在使用限制: 每月最多提供 2000 次代码补全和 50 次对话。

9、[OpenAI o3模型接近AGI](#)

2024.12.21, OpenAI 12天活动的最后一期, 主要介绍了新一代推理模型 o3 和 o3-mini。OpenAI声称, 至少在某些条件下, o3模型可以接近实现AGI。OpenAI计划在明年1月底前正式推出o3 mini, 之后推出完整版的o3。

o3模型在ARC-AGI基准上获得了破纪录的分数, ARC-AGI由Keras之父François Chollet开发, 主要是通过图形逻辑推理来测试模型的推理能力。以100%为最高分的ARC-AGI评估结果显示, 在低计算场景中, o3得分为75.7%, 而在高计算测试中, 它达到了87.5%。这标志着, o3的最佳成绩超过了标志着达到人类水平的门槛85%。作为对比, 目前开放的o1模型的得分仅在25%到32%之间。



三、技术文章

1、[基于 Ray 的大规模离线推理](#) (中文)

大模型离线推理，是指在具有数十亿或数千亿参数的大规模模型上进行分布式推理的过程。相较于常规模型推理，在模型切分、数据处理和数据流、提升 GPU 利用率方面面临了很大挑战。本次分享将介绍如何利用 Ray 及云原生优势助力大模型离线推理。

2、[复旦发布最佳RAG方案](#) (中文)

本文探讨了RAG技术的关键模块与优化方案，通过查询分类、片段划分、检索方法等模块的系统性研究，提出了性能与效率平衡的最佳实践，并展示了多模态扩展的潜力。

3、[不可变 Linux 发行版](#) (英文)

越来越多的开发者开始使用不可变的 Linux 发行版，即系统文件在运行期间不可更改。这样会让系统更稳定，也更容易迁移和回滚。

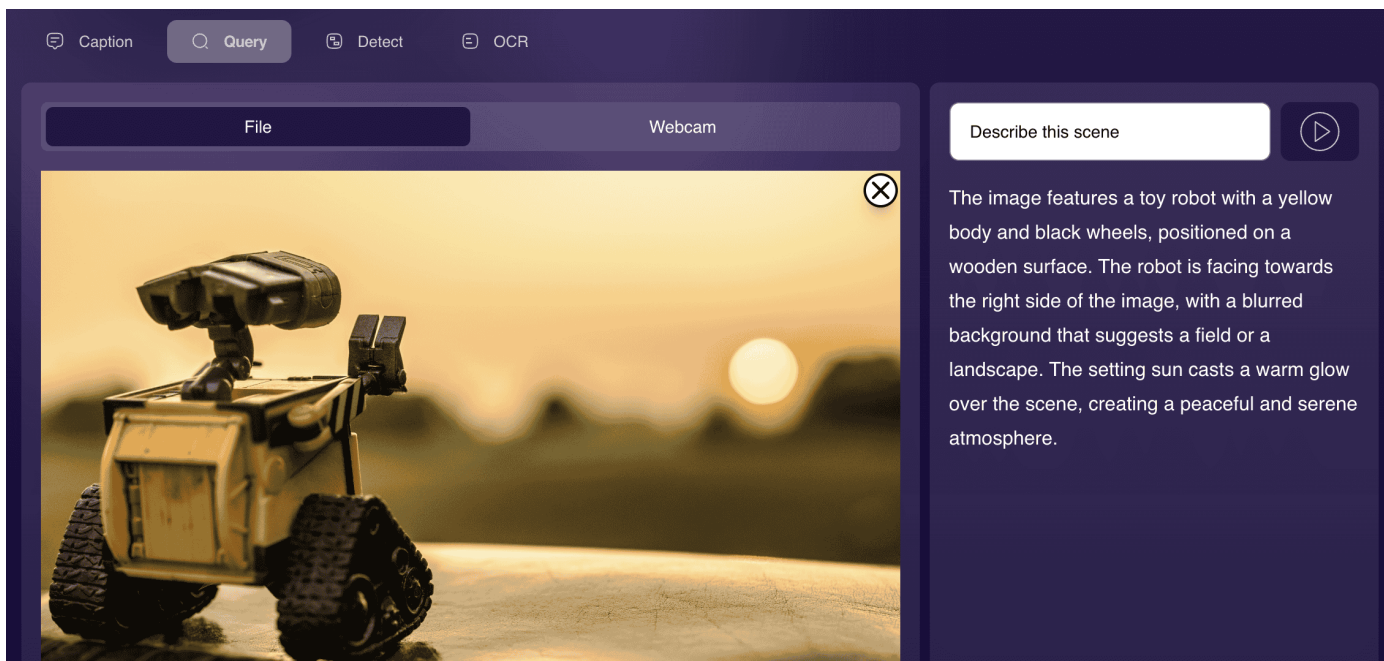
四、开源组件

1、[ragas](#)

Ragas是一个用于评估检索增强生成流程的框架，旨在帮助用户量化和改善RAG应用的性能。它可以集成到CI/CD中进行持续监控，并提供如ragas_score、context_precision等评估指标。

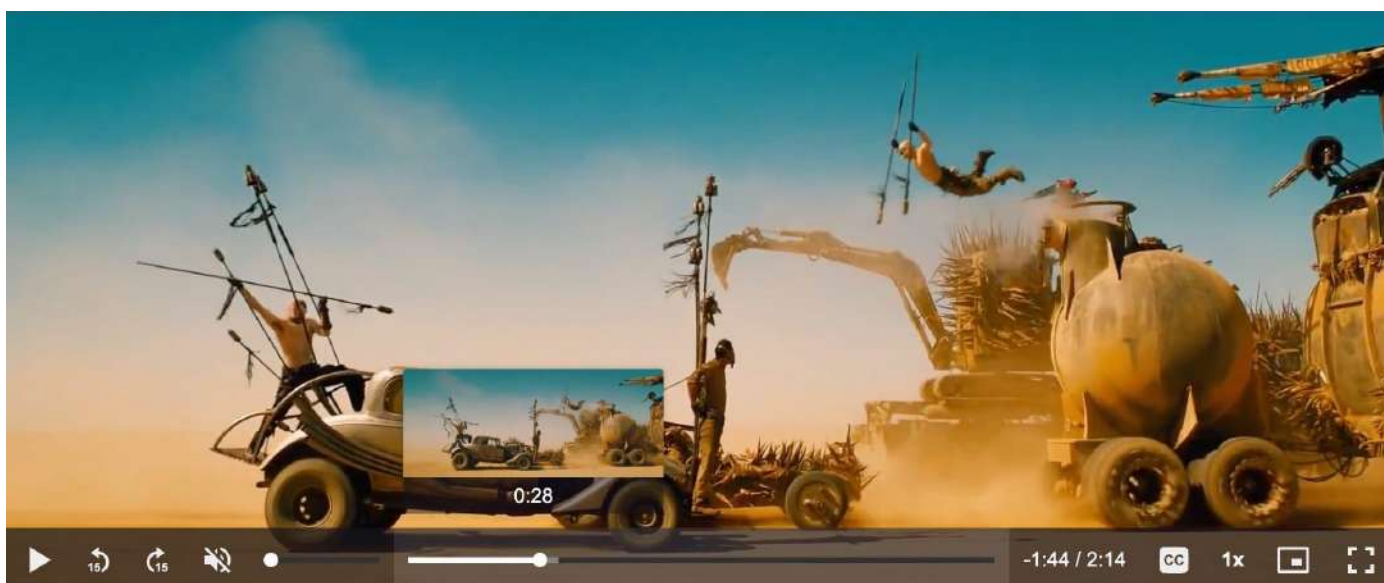
2、[moondream](#)

这是一个可在资源受限的设备上运行的小型视觉语言模型，它能够理解并生成与图像相关的自然语言描述，支持图像识别、生成描述和问答等功能。



3、[media-chrome](#)

这是一个用于定制网页音频和视频播放器界面的库，兼容各种 JavaScript 框架。它高度可定制，开发者可以轻松调整组件的外观和功能，支持字幕、投屏、快捷键、倍速、预览缩略图、移动端和静音按钮等功能。

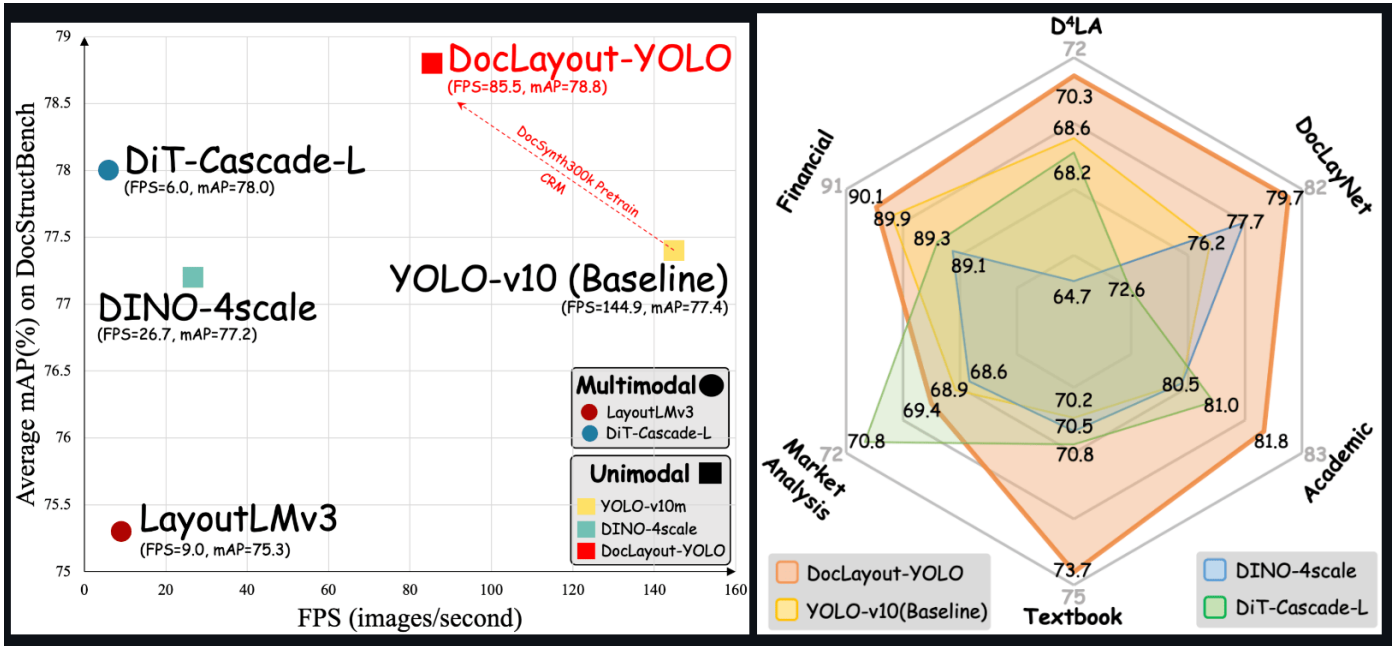


4、[Amphion](#)

Amphion是一个音频、音乐和语音生成工具包。它的目的是支持可重复的研究，并帮助初级研究人员和工程师进入音频、音乐和语音生成研究和开发领域。

5、[DocLayout-YOLO](#)

文档布局分析工具，通过多样化的合成数据和全局到局部的自适应感知增强文档布局分析，提升处理速度和准确性。[论文地址](#)

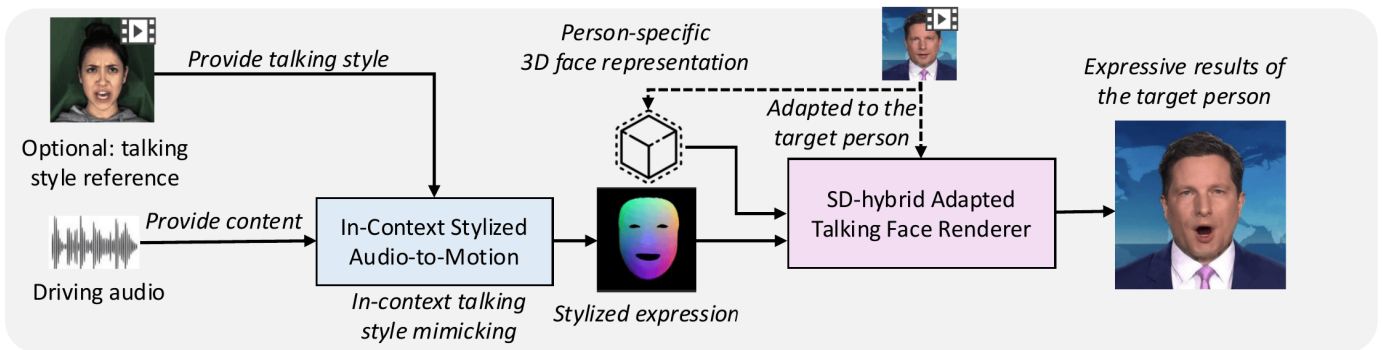


6、AdvancedLiteraryMachinery

是阿里巴巴读光OCR团队维护的开源项目，旨在构建具备高级智能的系统，赋予机器阅读、思考和创造的能力。

7、MimicTalk

MimicTalk是由浙江大学与字节跳动的联合研发团队开发的，它能够在短短15分钟内生成个性化和富有表现力的3D说话人脸视频。这项技术通过利用基于神经辐射场（NeRF）的泛化模型，结合静态-动态混合适应流程和上下文风格化的音频到运动模型，实现了对目标身份的静态外观和动态特征的快速学习与模仿，显著提高了视频质量、效率和表现力，超越了以往的个性化说话人脸生成方法。[论文地址](#)



8、PromptFix

PromptFix是一款基于扩散模型的图像处理工具，旨在根据用户的指令处理受损的图像并去除不需要的元素。它具有多种功能，包括图像上色、物体去除、去雾、去模糊、水印去除、雪花去除和低光增强等。该工具利用20步去噪过程，能够在修复图像缺陷的同时保留原始结构，并且适用于不同的纵横比。[官方介绍](#)

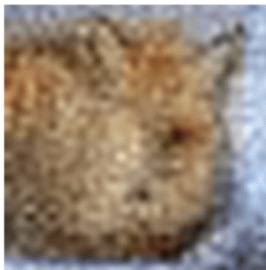
9、InstantIR

InstantIR是一种基于扩散模型的盲图像恢复方法，能够在测试时处理未知退化问题，提高模型的泛化能力。该技术通过动态调整生成条件，在推理过程中生成参考图像，从而提供稳健的生成条件。[官方介绍](#)



I. Photorealistic Image Restoration Low-Quality Input High-Quality Restoration

II. Text-Guided Edit



(a) Low-Quality Input



(b) w/o Prompt



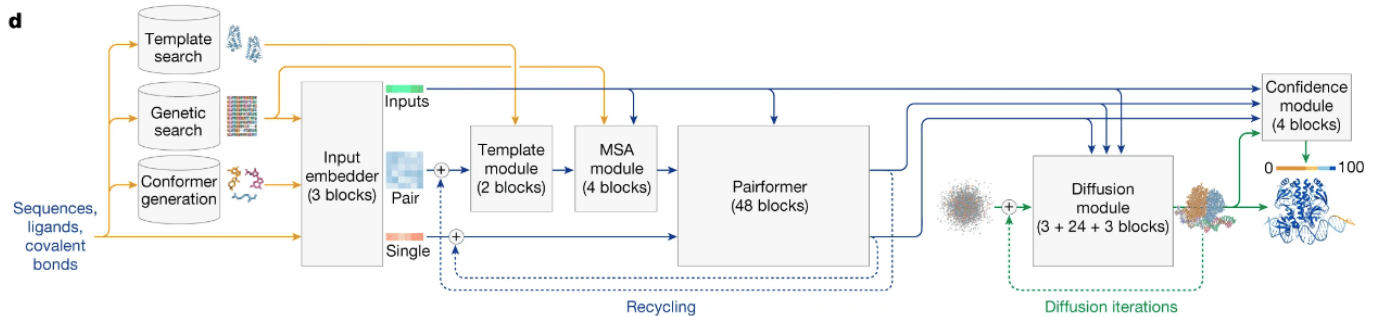
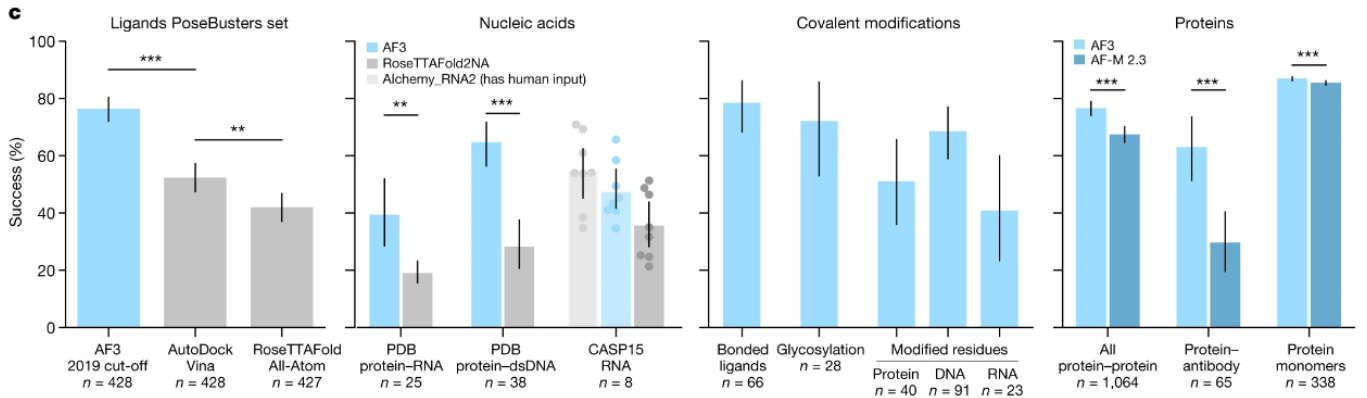
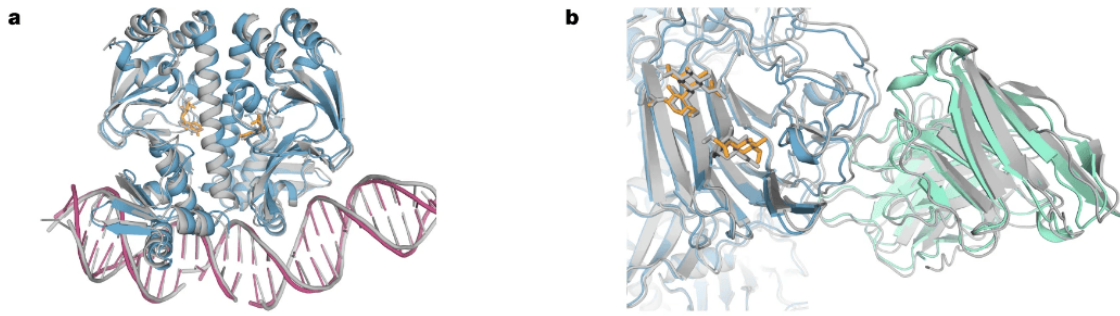
(c) Prompt: Fox



(d) Prompt: Cat

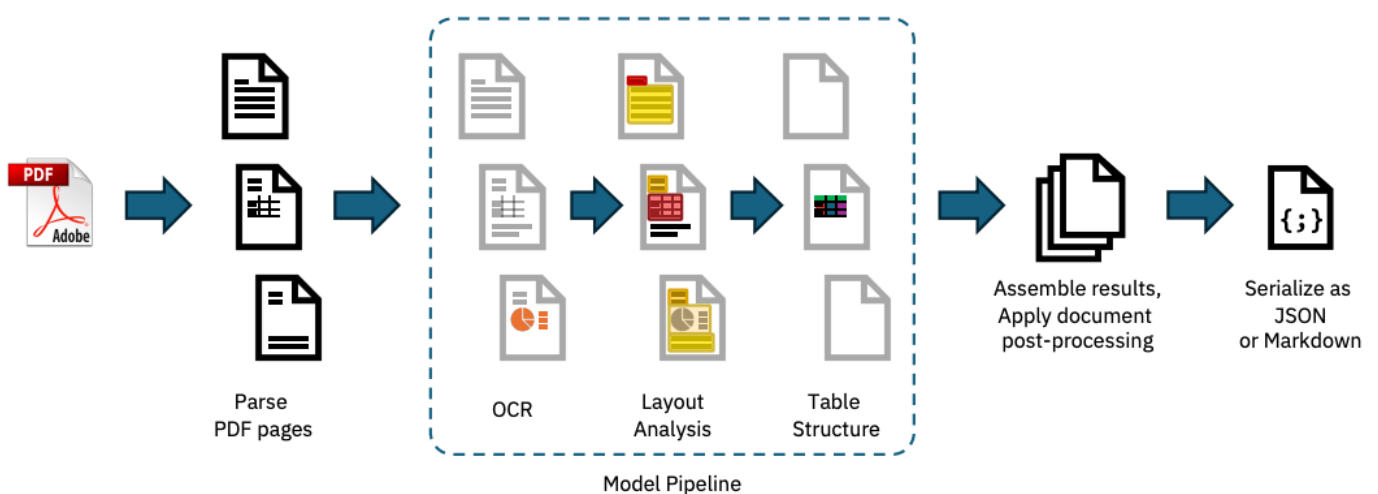
10、[AlphaFold3](#)

获得2024诺贝尔生物学奖的AlphaFold3蛋白质结构预测项目开源了（模型权重未直接放出，需要申请），它不仅能够预测蛋白质结构，还能预测其与DNA、RNA和其他蛋白质的相互作用，对于药物发现领域具有重大意义。[论文地址](#)、[在线体验](#)



11、[Docling](#)

Docling是一个文档解析和转换工具，能高效地将多种格式的文档（PDF、DOCX、PPTX、图片和HTML）解析导出为Markdown或JSON格式。Docling支持高级PDF理解、OCR功能，能与LlamaIndex和LangChain等工具集成，增强文档的检索和问答能力。[技术报告](#)



12、[PDFMathTranslate](#)

开源的命令行工具，使用 LLM 翻译 PDF 文件，可以完整保留排版。

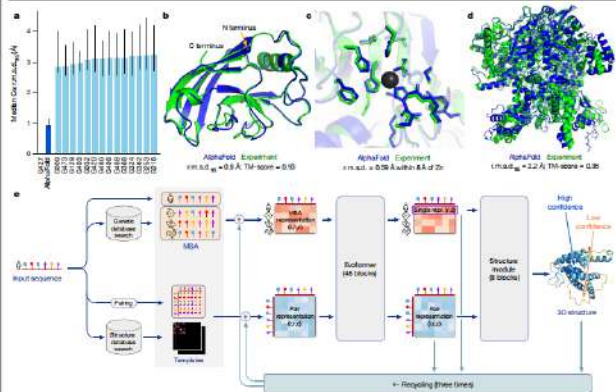


Fig. 1 | AlphaFold produces highly accurate structures. a, The performance of AlphaFold on the CASP14 dataset ($n = 87$ protein domains) relative to the top 15 methods (out of 440 entries), grouped by accuracy (responsibility numbers assigned to entries by CASP) data are median and the 95% confidence interval of median, estimated from 10,000 bootstrap samples. b, Our prediction of CASP14 target T1049 (PDB 6Y4I, blue) compared with the true (experiment) structure (grey). Four residues in the C-terminus of the crystal structure are factor-outside visual areas depicted. c, CASP14 target T1056 (PDB 6YJ1). An example for well-predicted zinc-binding site (AlphaFold has accurate side chains even though it does not explicitly predict the zinc). d, CASP target T1044 (PDB 6Y9J) = 2,150-residue single-chain was predicted with correct domain packing; the prediction was made after CASP using AlphaFold without intervention. e, Model architecture. Arrows show the information flow among the various components described in this paper. An arrow shape is shown in parentheses with the number of sequences (N_{seq}) in the main text; number of residues (N_{res}) in the main text; c, number of chains.

for the participating methods, and has long served as the gold-standard assessment for the accuracy of structure prediction^{23,24}. In CASP14, AlphaFold structures were vastly more accurate than competing methods. AlphaFold structures had a median backbone accuracy of 0.96 Å r.m.s.d., (C α root-mean-square deviation \pm 95% residue coverage) (95% confidence interval = 0.85–1.16 Å) whereas the next best performing method had a median backbone accuracy of 2.8 Å r.m.s.d., (95% confidence interval = 2.7–4.0 Å) (measured on CASP domain; see Fig. 1a for backbone accuracy and Supplementary Fig. 14 for all-atom accuracy). As a comparison point for this accuracy, the width of a carbon atom is approximately 1.4 Å. In addition to very accurate domain structures (Fig. 1b), AlphaFold is able to produce highly accurate side chains (Fig. 1c) when the backbone is highly accurate and considerably improves over template-based methods even when strong templates are available. The all-atom accuracy of AlphaFold was 1.5 Å r.m.s.d., (95% confidence interval = 1.2–1.6 Å) compared with the 3.5 Å r.m.s.d., (95% confidence interval = 3.1–4.2 Å) of the best alternative method. Our methods are scalable to very long proteins with accurate domains and domain packing (see Fig. 1d for the prediction of a 2,150-residue protein with no structural homologues). Finally, the model is able to provide precise, per-residue estimates of its reliability that should enable the confident use of these predictions. We demonstrate in Fig. 2a that the high accuracy that AlphaFold demonstrated in CASP14 extends to a large sample of recently released PDB structures; in this dataset, all structures were deposited in the PDB after our training data cut-off and are analysed as full chains (see Methods, Supplementary Fig. 15 and Supplementary Table 6 for more details). Furthermore, we observe high side-chain accuracy when the backbone predictions are accurate (Fig. 2b) and we show that our confidence measure, the predicted local-distance difference test (pLDDT), reliably predicts the C α local-distance difference test (lDDT-C α) accuracy of the corresponding prediction (Fig. 2c). We also find that the global superposition metric template modelling score (TM-score)²⁵ can be accurately estimated (Fig. 2d). Overall, these analyses validate that the high accuracy and reliability of AlphaFold on CASP14 proteins also transfers to an uncurated collection of recent PDB submissions, as would be expected (see Supplementary Methods L15 and Supplementary Fig. 11 for confirmation that this high accuracy extends to new folds).

The AlphaFold network
AlphaFold greatly improves the accuracy of structure prediction by incorporating novel neural network architectures and training procedures based on the evolutionary, physical and geometric constraints of protein structures. In particular, we demonstrate a new architecture to jointly represent sequence alignments (MSAs) and pairwise features, a new output representation and associated loss that enable accurate end-to-end structure prediction, a new equivalent attention

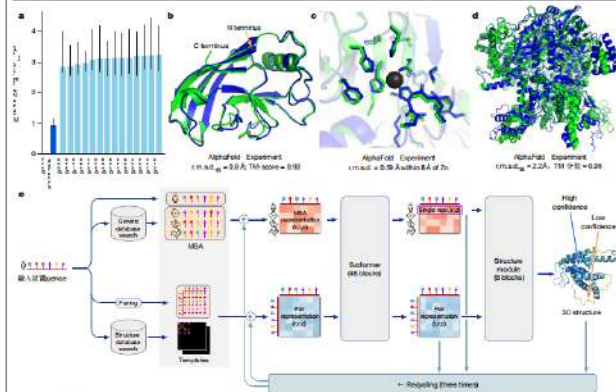


Fig. 1 | AlphaFold produces highly accurate structures. a, AlphaFold on CASP14 dataset ($n = 87$ protein domains) relative to the top 15 methods (out of 440 entries), grouped by accuracy (responsibility numbers assigned to entries by CASP) data are median and the 95% confidence interval of median, estimated from 10,000 bootstrap samples. b, Our prediction of CASP14 target T1049 (PDB 6Y4I, blue) compared with the true (experiment) structure (grey). Four residues in the C-terminus of the crystal structure are factor-outside visual areas depicted. c, CASP14 target T1056 (PDB 6YJ1). An example for well-predicted zinc-binding site (AlphaFold has accurate side chains even though it does not explicitly predict the zinc). d, CASP target T1044 (PDB 6Y9J) = 2,150-residue single-chain was predicted with correct domain packing; the prediction was made after CASP using AlphaFold without intervention. e, Model architecture. Arrows show the information flow among the various components described in this paper. An arrow shape is shown in parentheses with the number of sequences (N_{seq}) in the main text; number of residues (N_{res}) in the main text; c, number of chains.

network. In this dataset, all structures were deposited in the PDB after our training data cut-off and are analysed as full chains (see Methods, Supplementary Fig. 15 and Supplementary Table 6 for more details). Furthermore, we observe high side-chain accuracy when the backbone predictions are accurate (Fig. 2b) and we show that our confidence measure, the predicted local-distance difference test (pLDDT), reliably predicts the C α local-distance difference test (lDDT-C α) accuracy of the corresponding prediction (Fig. 2c). We also find that the global superposition metric template modelling score (TM-score)²⁵ can be accurately estimated (Fig. 2d). Overall, these analyses validate that the high accuracy and reliability of AlphaFold on CASP14 proteins also transfers to an uncurated collection of recent PDB submissions, as would be expected (see Supplementary Methods L15 and Supplementary Fig. 11 for confirmation that this high accuracy extends to new folds).

AlphaFold network
AlphaFold greatly improves the accuracy of structure prediction by incorporating novel neural network architectures and training procedures based on the evolutionary, physical and geometric constraints of protein structures. In particular, we demonstrate a new architecture to jointly represent sequence alignments (MSAs) and pairwise features, a new output representation and associated loss that enable accurate end-to-end structure prediction, a new equivalent attention

13、slugify

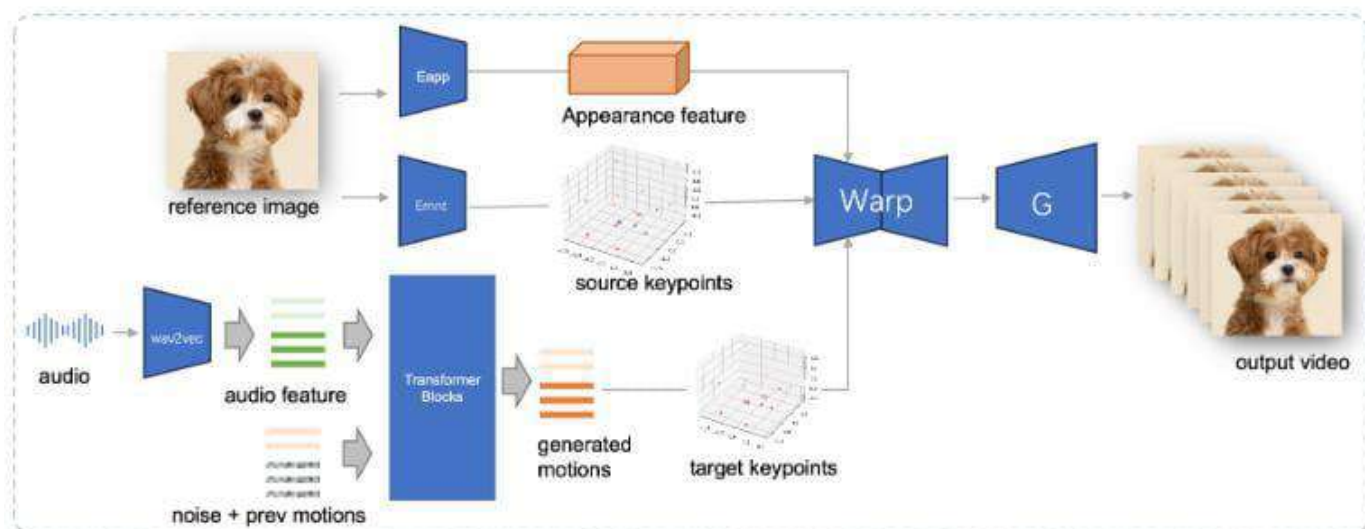
该项目是用于将字符串转换为适合在 URL 中使用的格式，输出由小写字母、数字和短横线组成的字符串，不含空格和特殊字符，这种格式有助于搜索引擎优化（SEO）。

14、imagehash

该项目的算法不同于传统的加密哈希算法（如 MD5、SHA-1），它专注于图像内容的相似度分析，对有细微不同的图片可生成相似的哈希值，用于计算图片相似度，支持平均哈希、感知哈希、差分哈希等算法，适用于快速识别版权图片等场景。

15、JoyVASA

JoyVASA是一种基于扩散模型的音频驱动面部动画技术，它能够生成逼真的面部动态和头部运动。



16、[pyarmor](#)

这是一个用于对 Python 脚本进行混淆处理的命令行工具，仅需一条命令即可完成加密操作。它提供丰富的加密选项，用来平衡安全与性能，支持将加密后的脚本绑定到特定机器、设置加密有效期和 Themida 保护等功能。

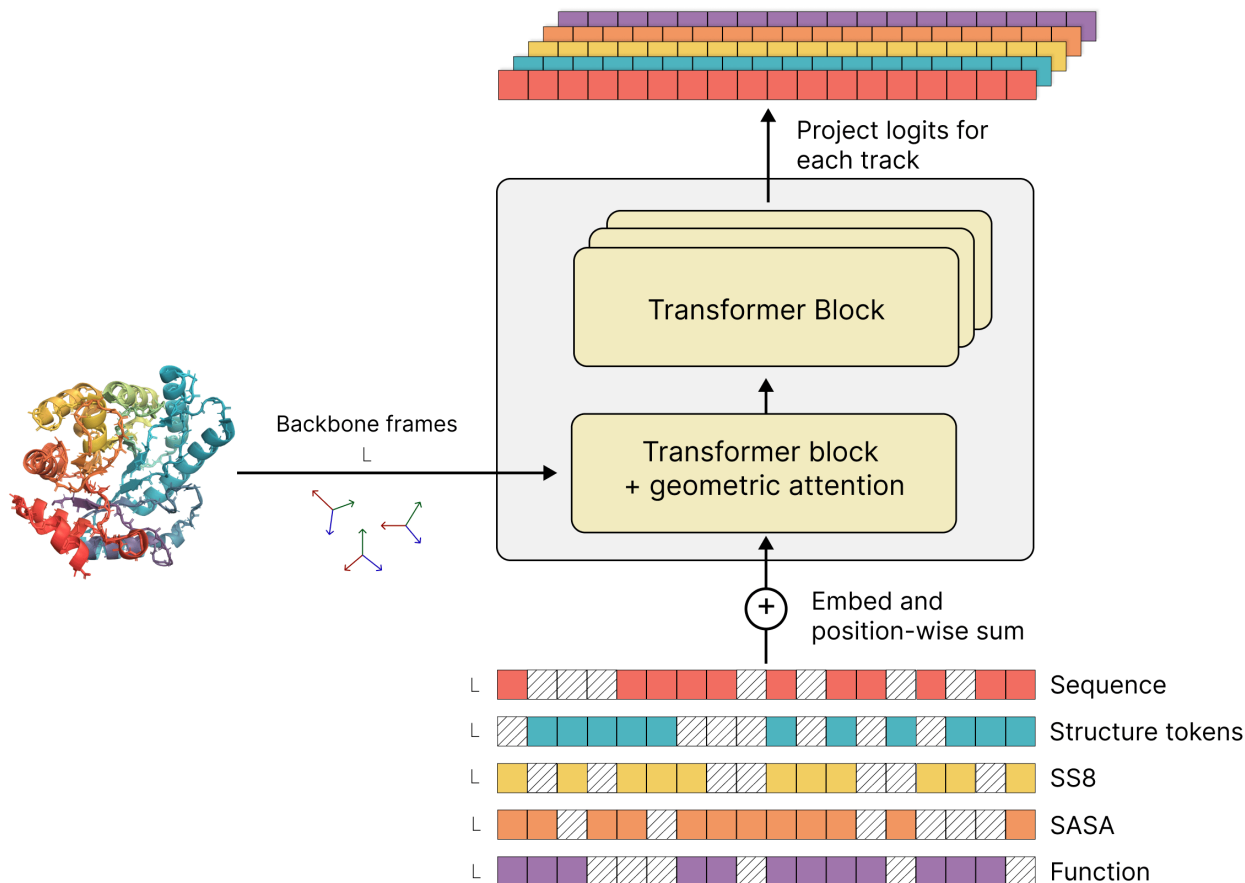
17、[TensorRT-YOLO](#)

这是一款专为 NVIDIA 设备优化的 YOLO 部署工具。它通过集成 TensorRT 插件和 CUDA 技术，提供 C++ 和 Python API，显著提升了推理速度和易用性，支持多种 YOLO 版本，适用于目标检测、实例分割、姿态识别、旋转目标检测和视频分析等多种场景。

18、[ESM](#)

ESM是一个用于蛋白质生成和表示的开源模型库，专注于生物学底层的表征。

- 提供多种参数规模的 ESM C 模型，具有更高效能和更低内存需求。
- ESM C 在性能上超越了之前版本 ESM2，尤其是 6B 参数模型设立了新的基准。
- 支持通过 GitHub、Forge API 和 AWS SageMaker 等多种方式使用模型。



19、[MegaParse](#)

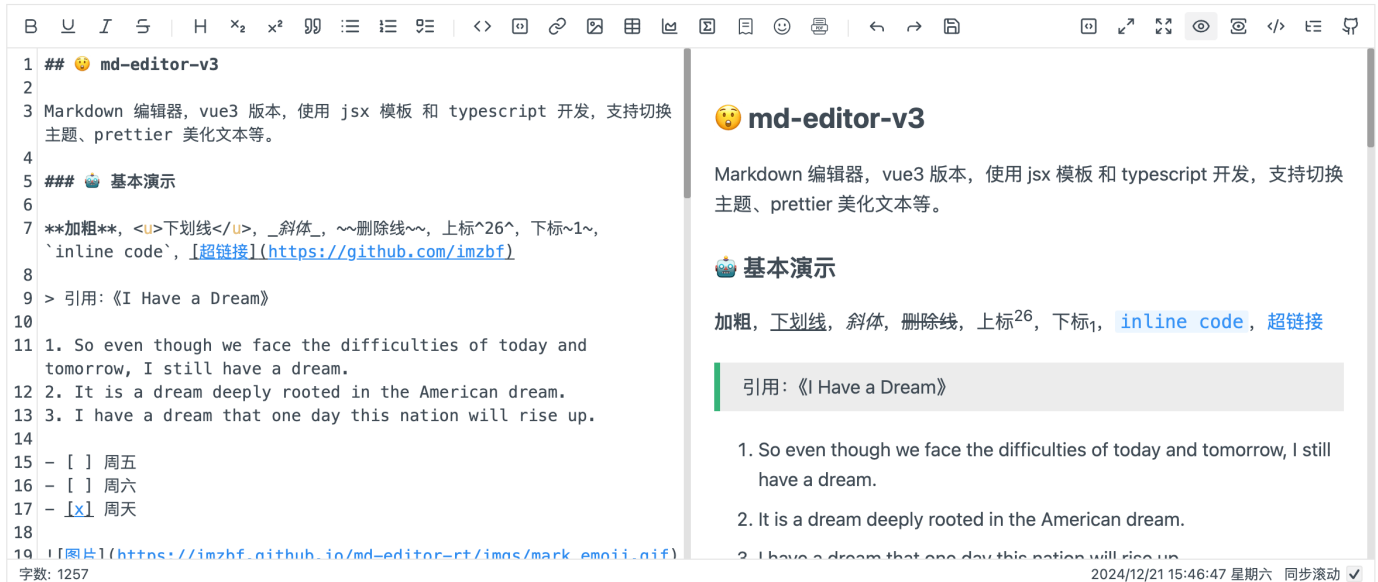
MegaParse是一个强大且通用的解析器，可以轻松处理各种类型的文档，包括文本、PDF、PowerPoint 演示文稿、Word 文档等。它旨在在解析过程中尽可能减少信息丢失。

20、[MarkItDown](#)

Markitdown是微软最新推出的一款在线文档转换工具，其核心功能是将多种格式的文件（包括PDF、PPT、Word、Excel、HTML等）转换为Markdown格式，可用于构建RAG系统。

21、[md-editor-v3](#)

md-editor-v3是一个基于Vue 3和TSX开发的Markdown编辑器组件，支持实时预览和丰富的Markdown功能。[官方文档](#)

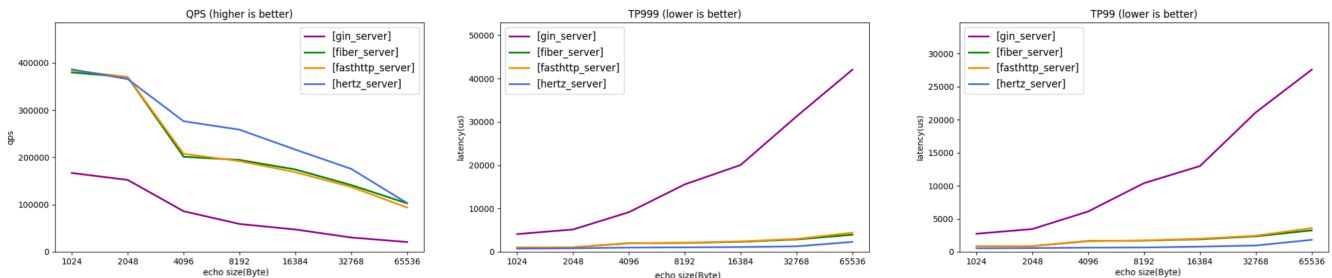


22、[AntV G6](#)

AntV G6是一个简单、易用、完备的图可视化引擎，它在高定制能力的基础上，提供了一系列设计优雅、便于使用的图可视化解决方案。能帮助开发者搭建属于自己的图可视化、图分析、或图编辑器应用。[官方文档](#)

23、[hertz](#)

该项目是由 CloudWeGo 团队开发的高性能、易扩展的 Go 语言 Web 框架，可用于构建 RESTful API、微服务以及需要高吞吐量的 Web 应用。在字节跳动内部，Hertz 已被广泛应用，表现稳定可靠。



24、[Loguru](#)

Loguru是一个功能强大且易于使用的开源日志记录库。它建立在 Python 标准库中的 logging 模块之上，并提供了更加简洁直观的接口。使用 Loguru，我们可以轻松地记录不同级别的日志消息，并根据需求输出到终端、文件或其他目标。

25、[zap](#)

该项目是 Uber 开发的一款高性能日志库，专为那些需要快速、结构化日志记录的场景而设计。与其他 Go 语言的日志库相比，zap 的性能更为优越，尤其是在需要频繁记录日志的高并发环境中。

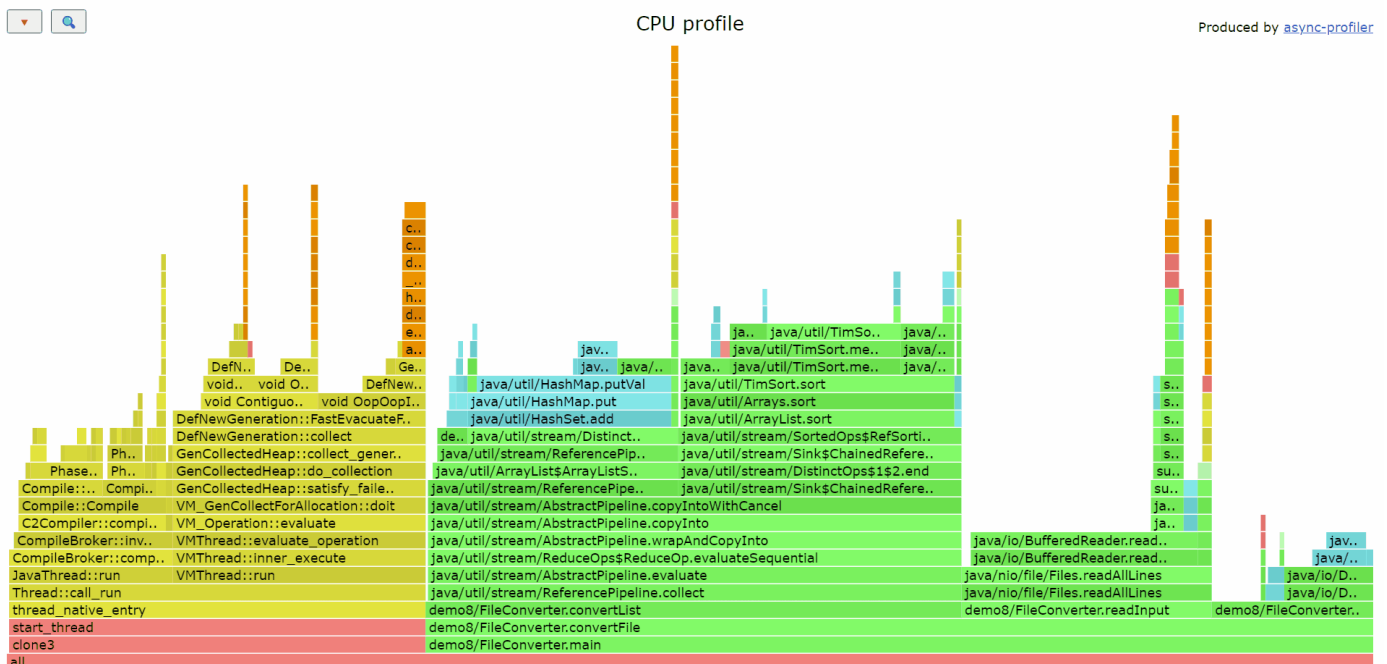
Package	Time	Time % to zap	Objects Allocated
⚡ zap	656 ns/op	+0%	5 allocs/op
⚡ zap (sugared)	935 ns/op	+43%	10 allocs/op
zerolog	380 ns/op	-42%	1 allocs/op
go-kit	2249 ns/op	+243%	57 allocs/op
slog (LogAttrs)	2479 ns/op	+278%	40 allocs/op
slog	2481 ns/op	+278%	42 allocs/op
apex/log	9591 ns/op	+1362%	63 allocs/op
log15	11393 ns/op	+1637%	75 allocs/op
logrus	11654 ns/op	+1677%	79 allocs/op

26、[letta](#)

用于构建具有记忆功能的 LLM 应用的 Python 框架，支持创建拥有长期记忆和持久状态的 Agent，并能够集成多种 LLM API 服务。

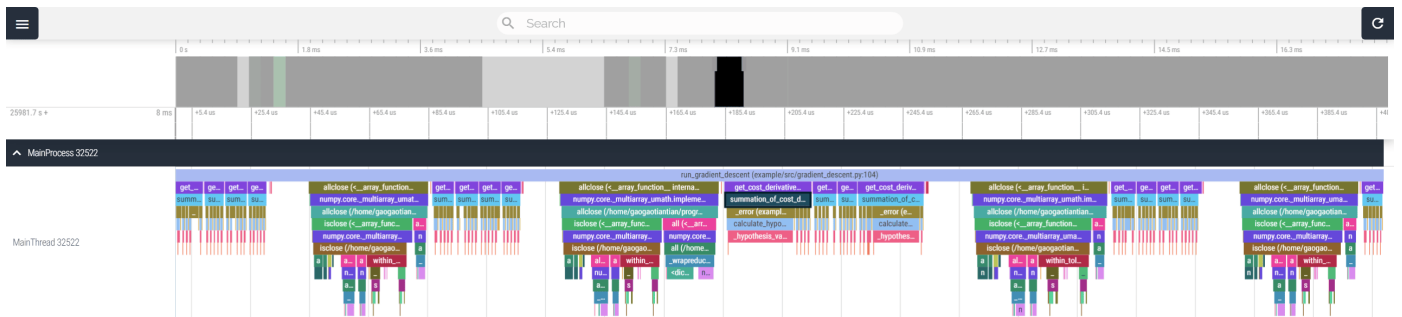
27、[async-profiler](#)

这是一个低开销的 Java 采样分析器，可用于对 Java 应用进行性能分析。它能够帮助开发者追踪 CPU 消耗的热点代码路径、定位内存泄漏，以及分析线程锁争用和线程间的同步等问题。



28、viztracer

一款高效、易用的 Python 程序性能分析工具，帮助开发者调试代码和分析性能瓶颈。它针对性能追踪做了优化，尽可能减少对被分析程序的性能影响，支持时间线追踪、多进程分析和模块化追踪等功能，并生成直观、交互式的分析报告。



```
Current Selection
Slice Details
Name      summation_of_cost_derivative (example/src/gradient_descent.py:75)
Category  FEE
Start time 8ms 192us 700ns
Duration  29us
type      internal_slice
track_id  0
depth    2
stack_id  2372463328701694
parent_stack_id 2597963731413117
parent_id 6802
arg_set_id 0

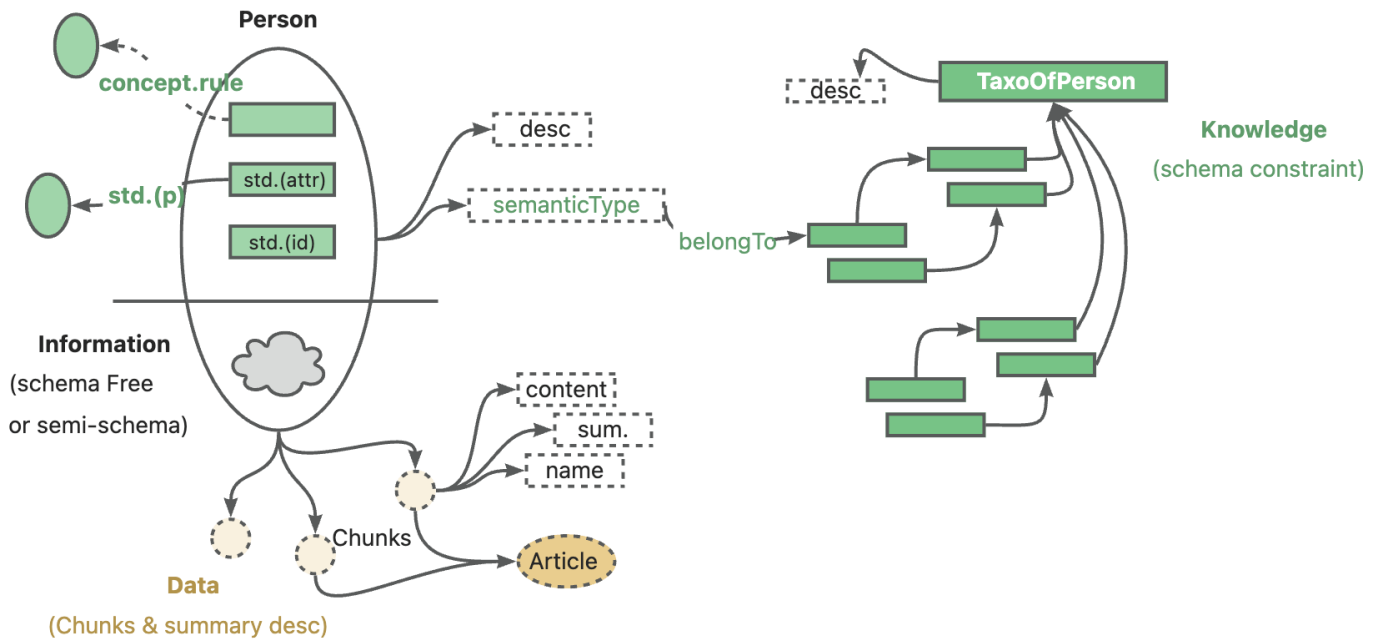
75 def summation_of_cost_derivative(index, end-m):
76     """
77     Calculates the sum of cost function derivative
78     :param index: index wrt derivative is being calculated
79     :param end: value where summation ends, default is n, number of examples
80     :return: Returns the summation of cost derivative
81     Note: If index is -1, this means we are calculating summation wrt to biased
82     parameter.
83     """
84     summation_value = 0
85     for i in range(end):
86         if index == -1:
87             summation_value += _error(i)
88         else:
89             summation_value += _error(i) * train_data[i][0][index]
90     return summation_value
91
92
93 def get_cost_derivative(index):
94     """
95     :param index: index of the parameter vector wrt to derivative is to be calculated
96     :return: derivative wrt to that index
```

29、kuberaay

KubeRay是由字节跳动技术团队牵头，由蚂蚁金服、微软等公司共同参与建设的开源Ray部署集成工具集，目前已成为在 Kubernetes集群上部署Ray应用的事实标准。KubeRay提供了一个自定义资源定义，允许用户通过简单的YAML文件来定义Ray集群的配置。此外，它还提供了一组控制器，用于自动管理和扩展Ray集群。

30、KAG

KAG 是基于 OpenSPG 引擎和大型语言模型的逻辑推理问答框架，用于构建垂直领域知识库的逻辑推理问答解决方案。KAG 可以有效克服传统 RAG 向量相似度计算的歧义性和 OpenIE 引入的 GraphRAG 的噪声问题。KAG 支持逻辑推理、多跳事实问答等，并且明显优于目前的 SOTA 方法。



五、工具软件

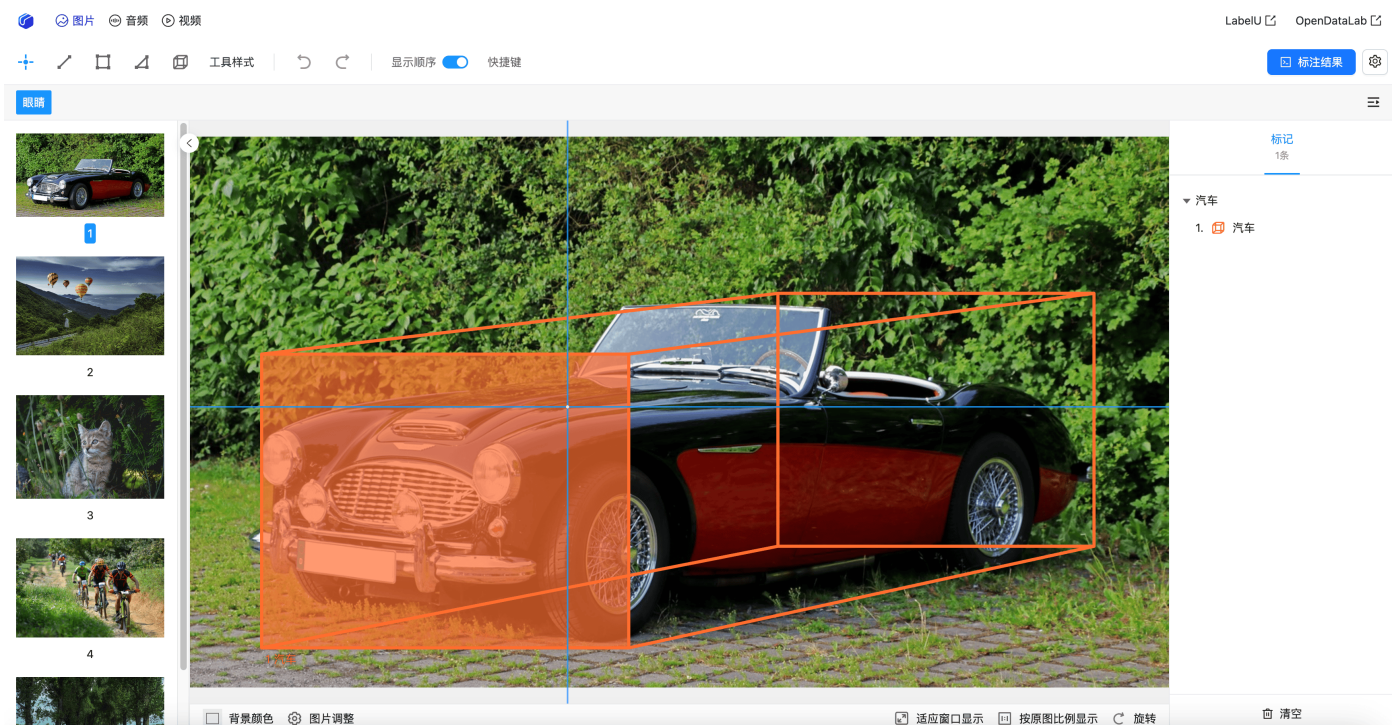
1、[Rufus](#) (开源)

Rufus是一款老牌的USB启动盘制作工具，可以将Windows、Linux等系统的ISO格式的镜像制作成USB启动盘。

2、[LabelU](#) (开源)

LabelU是一款综合性的数据标注平台，专为处理多模态数据而设计。该平台旨在通过提供丰富的标注工具和高效的工作流程，帮助用户更轻松地处理图像、视频和音频数据的标注任务，满足各种复杂的数据分析和模型训练需求。

[在线体验](#)

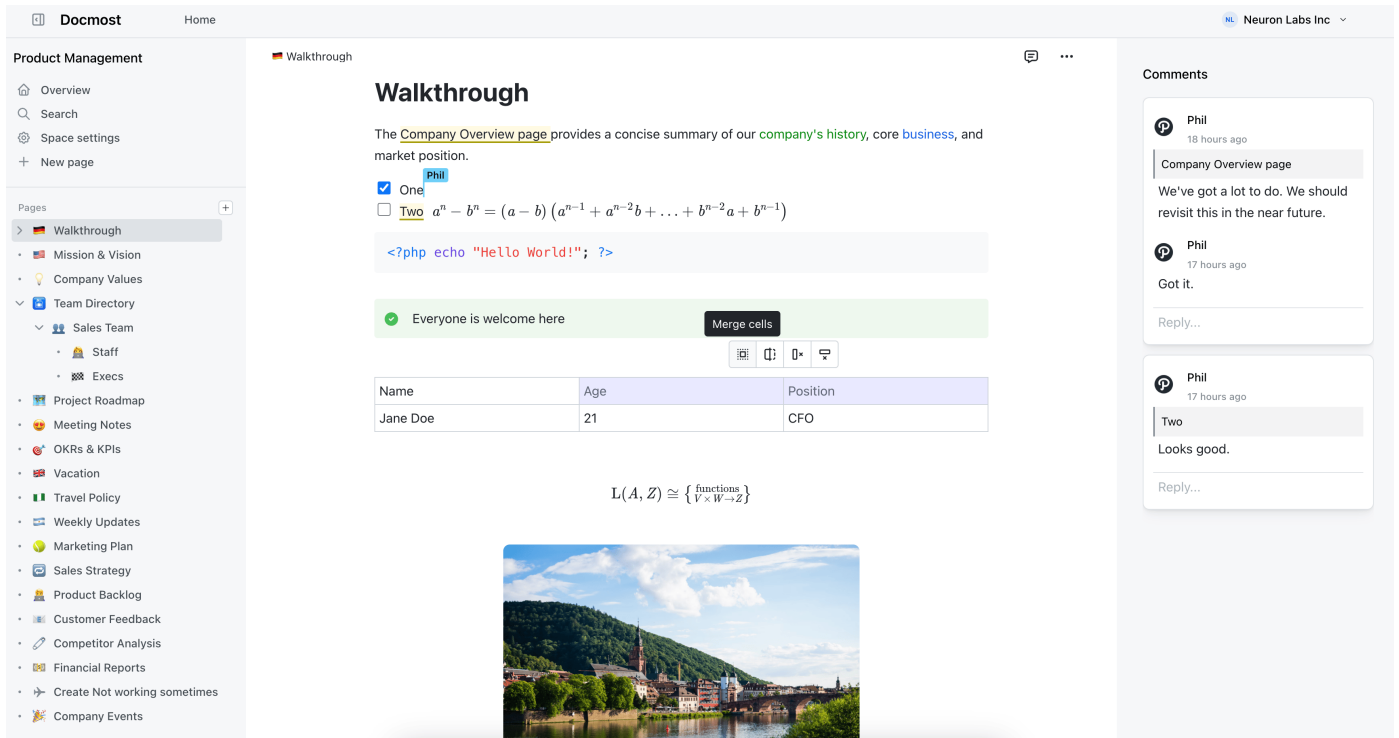


3、[Cursor](#) (免费)

Cursor是一款基于人工智能技术的代码生成工具，它能够根据开发者的需求和上下文，智能地生成高质量的代码片段。无论是初学者还是资深开发者，都可以通过Cursor快速完成复杂的编程任务，提高开发效率和质量。

4、[Docmost](#) (开源)

一个基于浏览器的维基和文档工具，Confluence 和 Notion 替代品。



5、[notifier](#) (开源)

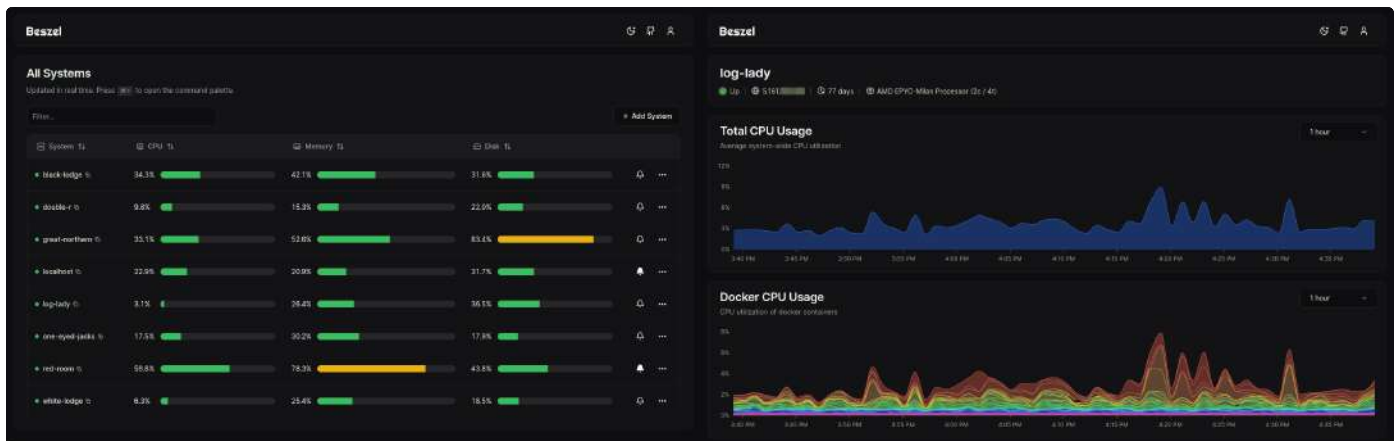
一个 Cloudflare Worker 应用，监听 GitHub 仓库的各种事件，自动发送信息到 Telegram。

6、[wechat-article-exporter](#) (开源)

开源的微信公众号文章在线导出工具，可用来批量下载某个公众号的文章。

7、[beszel](#) (开源)

量级高颜值的 Docker 监控平台。这是一个轻量级的服务器监控平台，包括 Docker 统计、历史数据和警报功能。它拥有友好的 Web 界面，配置简单、开箱即用，支持自动备份、多用户、OAuth 认证和 API 访问等功能。



8、[gophish](#) (开源)

该项目提供了一个开箱即用的网络钓鱼平台，可用于模拟钓鱼攻击。它拥有友好的 Web 管理后台，支持邮件模板、批量发送邮件、网站克隆和数据可视化，适用于企业安全培训和渗透测试等场景。

9、[kubernetes-goat](#) (开源)

该项目是用于构建漏洞百出、易受攻击的集群环境，让开发者可以在真实场景中学习 K8s 攻击和防御技巧。

10、[GoGoGo](#) (开源)

该项目是一个基于 Android 调试 API 和百度地图实现的虚拟定位工具，无需 ROOT 权限即可修改地理位置。它支持位置搜索和手动输入坐标，并提供了一个可自由移动的摇杆来模拟位移。



11、[oha](#) (开源)

这是一个用 Rust 开发的 HTTP 请求压测工具，它操作简单、带 TUI 动画界面，支持生成请求延迟、吞吐量等指标的报告，以及动态 URL 和更灵活请求间隔等功能。

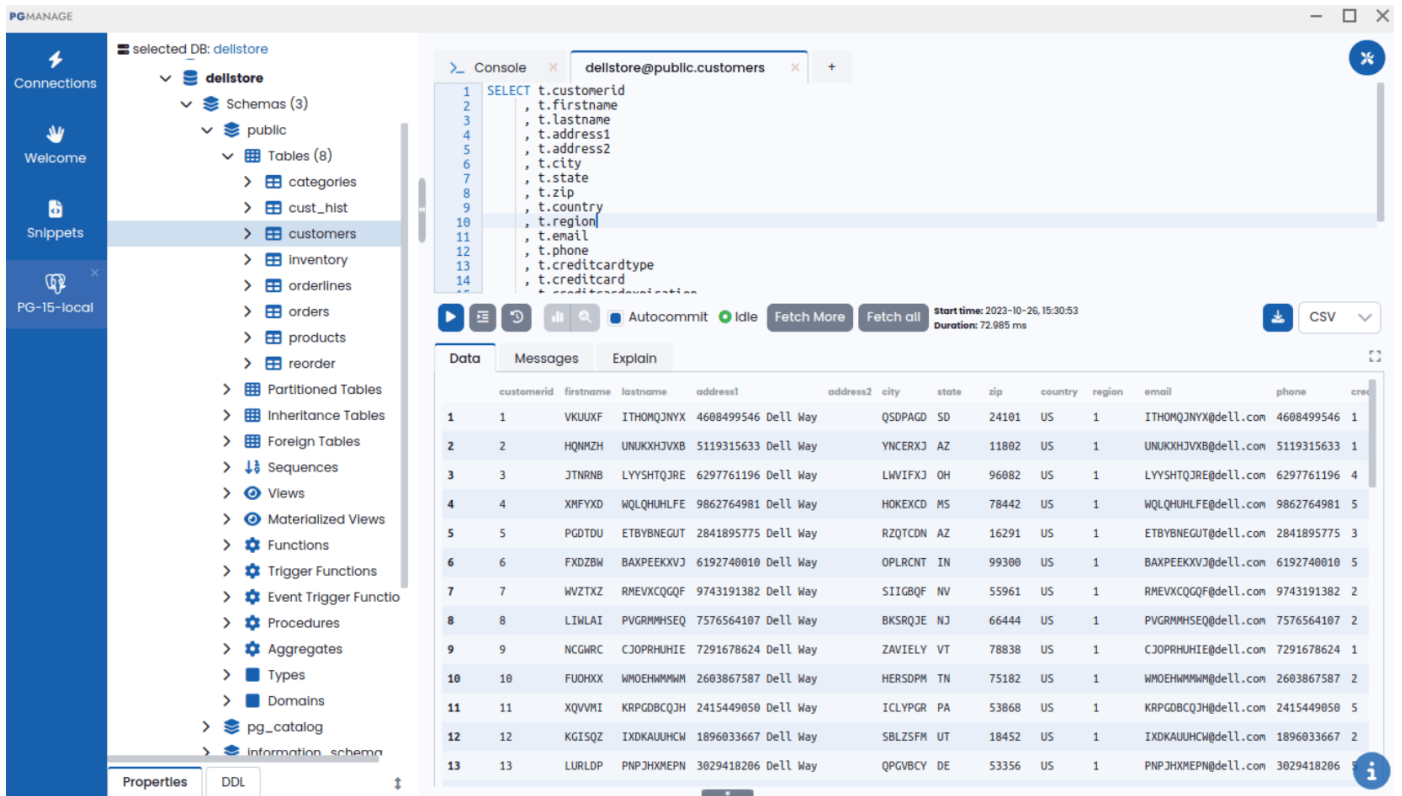
12、[BilibiliSponsorBlock](#) (开源)

这是一款能够自动跳过 B 站视频中恰饭片段和开场、结尾动画的浏览器插件，所有标注数据均由网友贡献。

[Chrome插件商店安装](#)

13、[PgManage](#) (开源)

一个跨平台的 Postgres 数据库的桌面客户端，提供图形界面。



14、[dockerc](#) (开源)

该项目能将 Docker 镜像转化为二进制可执行文件，无需配置 Docker 环境或安装依赖，简化了软件的分发和运行流程。

15、[Zed](#) (开源)

Zed 是一款支持多人协作的代码编辑器，底层采用 Rust，主打“高性能”，充分利用多核 CPU 和 GPU，追求极致性能，速度极快。

16、[302.AI](#) (开源)

302.AI 是一个“AI应用超市”，官方把它的应都开源出来了，你可以基于Next.js在这基础上二次开发自己的AI应用。[官方体验](#)

基于模型深度开发应用

不仅仅是对话机器人，我们提供真正的AI应用



17、[Electronic-Component-Sorter](#) (开源)

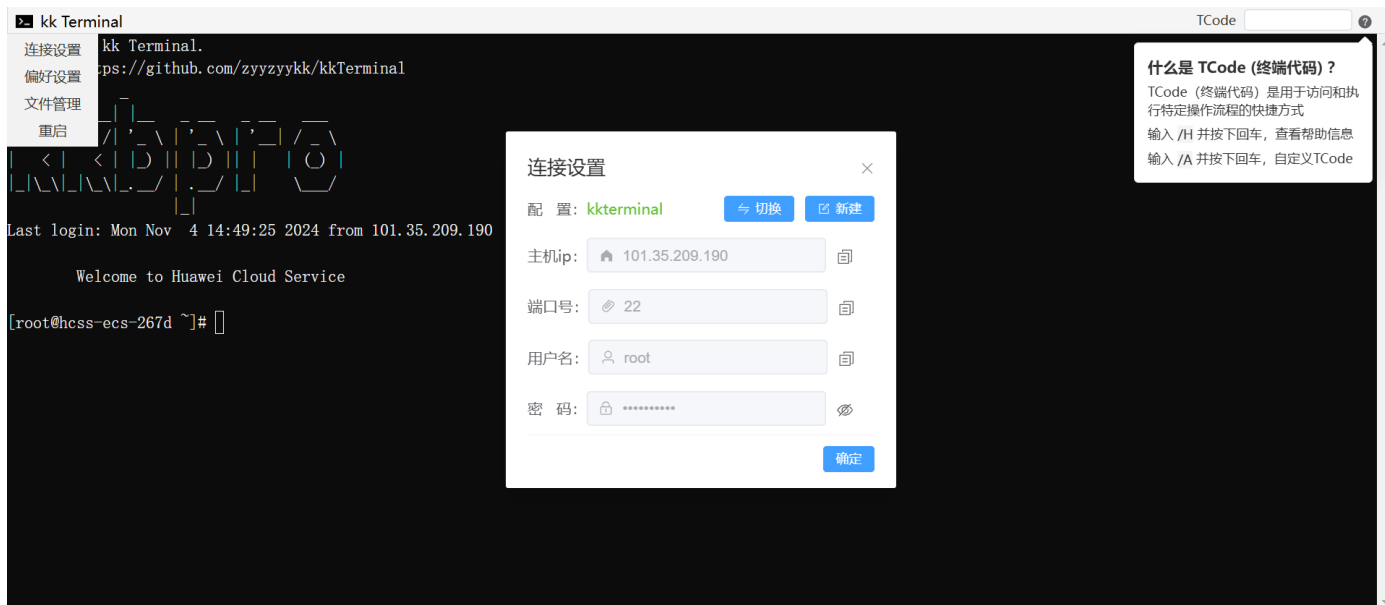
一个本机的AI工具，用摄像头识别电子元件。

18、[Exo](#) (开源)

Exo是一款实验性软件，允许用户通过家中的常见设备（如iPhone、iPad、Android、Mac、Linux等）组建自己的AI集群，进而将这些设备统合为一个强大的GPU。

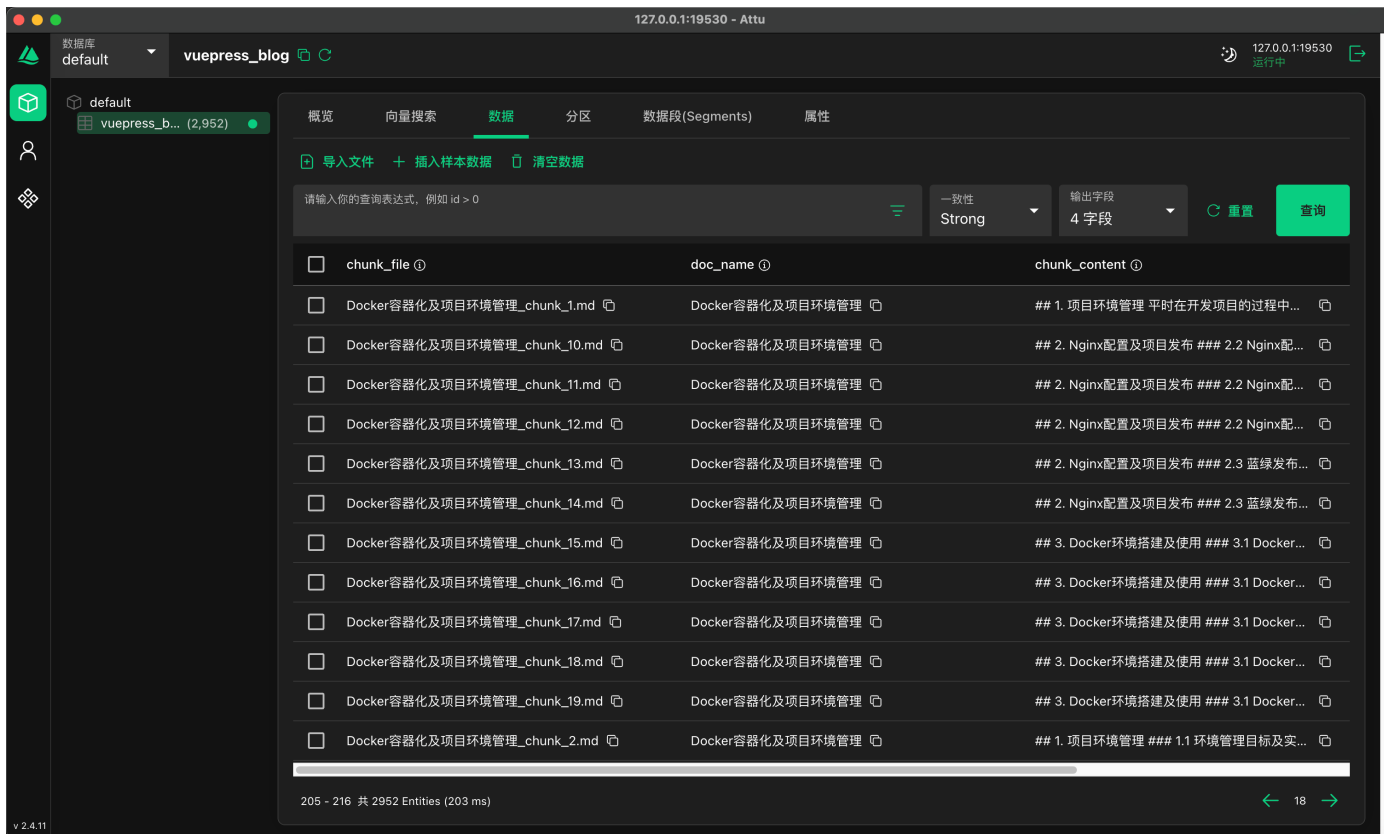
19、[kkTerminal](#) (开源)

一个开源的网页终端，在网页上通过SSH连入远程主机。



20、[Attu](#) (开源)

一款专为Milvus向量数据库打造的开源数据库管理工具，提供了便捷的图形化界面，极大地简化了对Milvus数据库的操作与管理流程。



21、[WebVM](#) (开源)

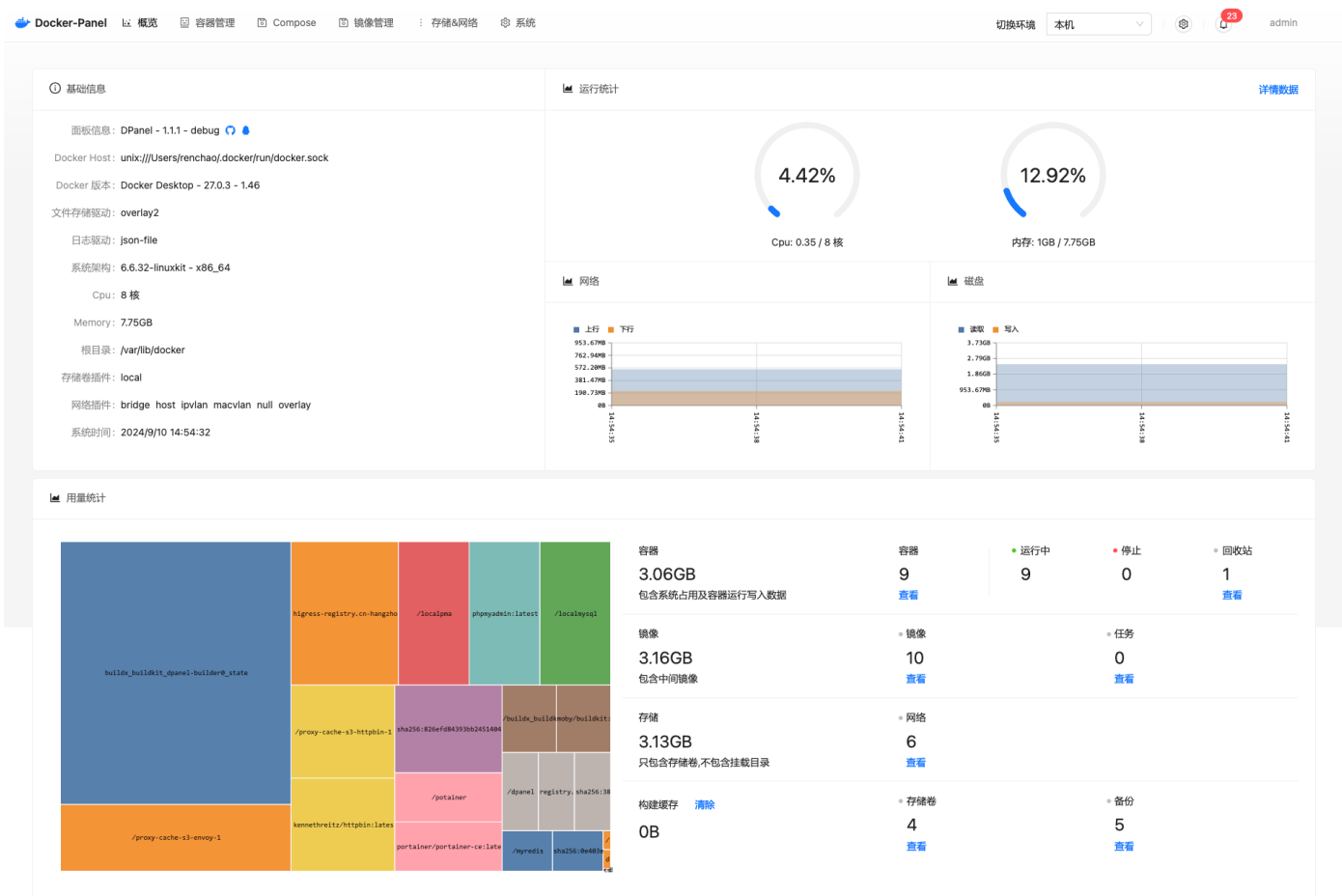
WebVM是一个创新的开源项目，它将Linux虚拟机直接搬到了你的浏览器中，无需服务器，即可体验全功能的Linux环境。这个项目旨在提供一个可定制的、安全的、沙箱化的客户端执行平台，让开发者和爱好者能在任何设备上轻松操作Linux。

22、[Garak](#) (开源)

Garak 是一款专用于扫描大语言模型（LLM）漏洞的工具，能够高效识别基于语言模型的系统中的潜在弱点和不良行为。它主要检测幻觉生成、数据泄露、提示注入、错误信息、毒性生成以及越狱等问题，并生成详细报告，概述模型的优势和改进建议。

23、[dpanel](#) (开源)

这是一款专为国内用户设计的 Docker 可视化管理面板，采用全中文界面。它安装简单且资源占用低，运行在容器内部对宿主机无侵入，支持容器管理、镜像管理、文件管理以及 Compose 管理等功能。



24、[kyanos](#) (开源)

这是一个基于 eBPF 的网络问题分析工具，能够实时监控和分析 HTTP、Redis 和 MySQL 请求。它支持强大的流量过滤功能，可根据进程、容器、协议信息和耗时等条件进行精确过滤，并提供多维度聚合抓取的数据包信息，适用于排查远程服务慢查询等问题。

25、[minisign](#) (开源)

一个开箱即用的文件数字签名与验证工具，只需要简单的命令即可生成和验证文件签名。它基于 Ed25519 公钥签名系统，提供可靠的文件完整性验证功能，适用于软件分发和文件共享等场景。

26、[Lute](#) (开源)

这是一个用 Go 语言编写的 Markdown 引擎，实现了最新的 GFM/CM 规范。它是将 Markdown 文本转换成一个抽象语法树 (AST)，无需正则表达式解析速度更快，支持 GFM/CM 规范、内置代码高亮、术语修正、格式化（中英文间自动插入空格）和 Emoji 解析等功能。

Markdown 原文:

在github上做开源项目是一件很开心的事情, 请不要把Github拼写成`github`哦!

特别是简历中千万不要出现这样的情况:

> 熟练使用JAVA、Javascript、GIT, 对android、ios开发有一定了解, 熟练使用Mysql、postgresql数据库。

修正后:

在 GitHub 上做开源项目是一件很开心的事情, 请不要把 GitHub 拼写成`github`哦!

特别是简历中千万不要出现这样的情况:

> 熟练使用 Java、JavaScript、Git, 对 Android、iOS 开发有一定了解, 熟练使用 MySQL、PostgreSQL 数据库。

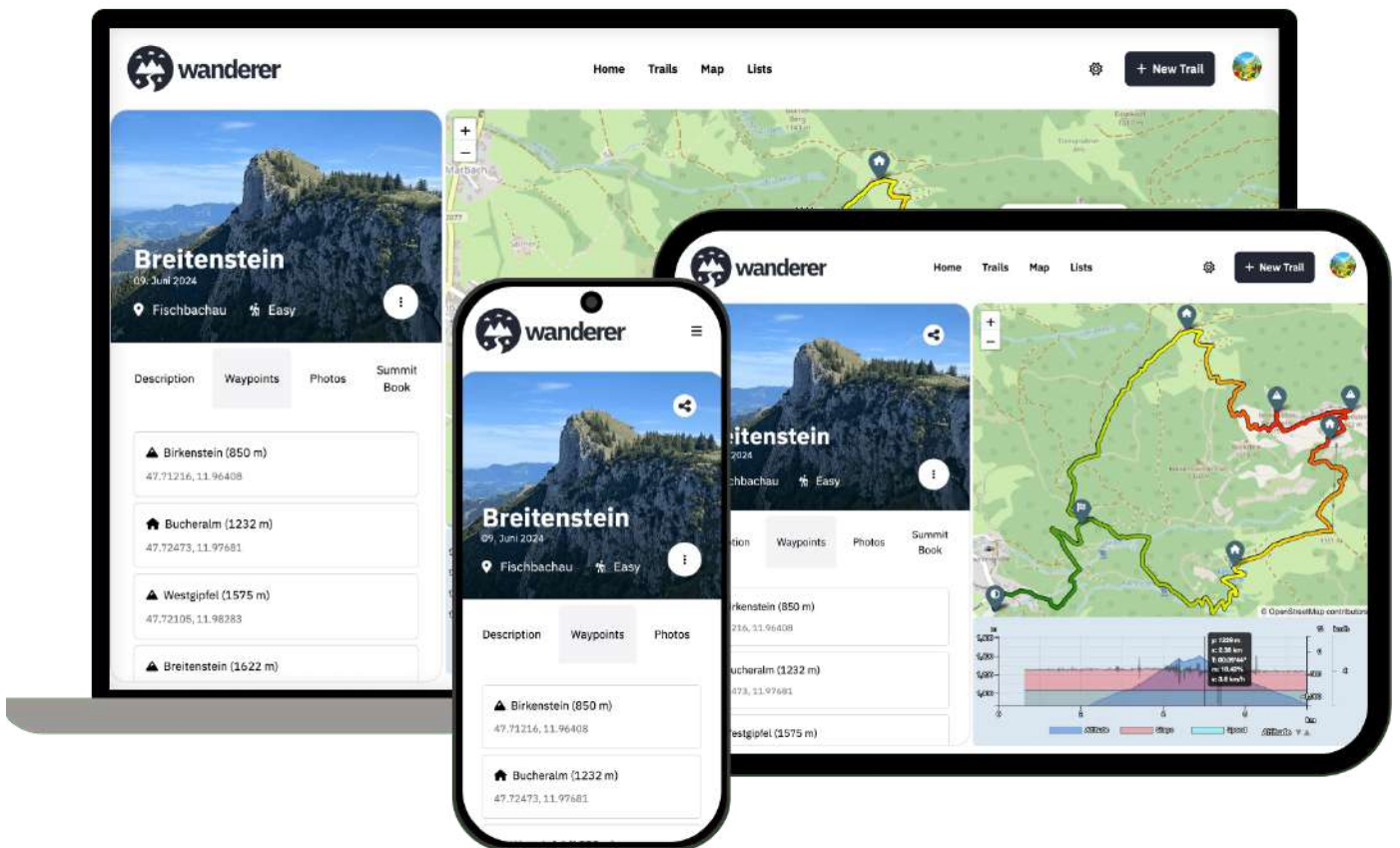


27、[Starlight](#) (开源)

该项目是基于 Astro 框架打造的文档主题, 可用于快速搭建和部署文档网站。它界面美观、开箱即用、访问速度快, 支持网站导航、搜索、国际化、SEO 和各种插件。

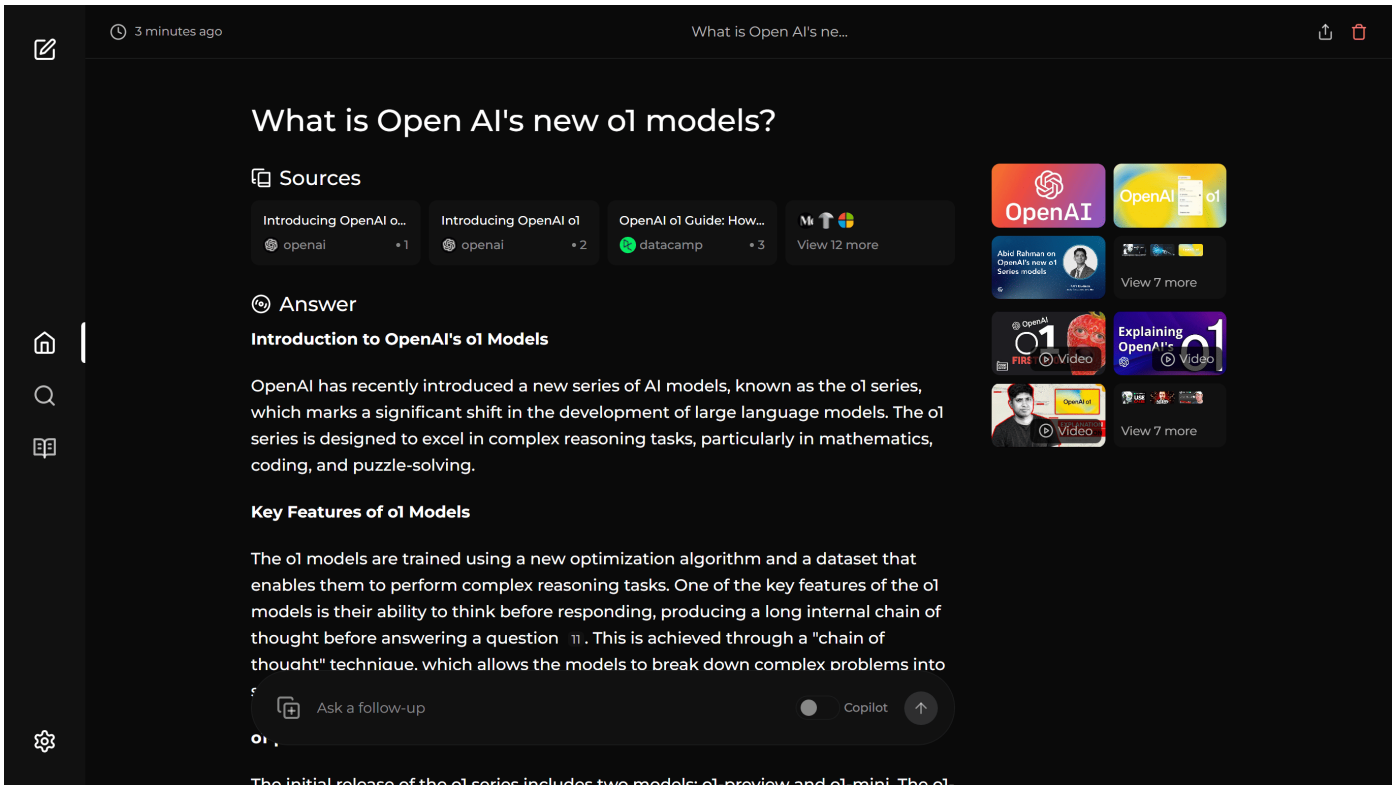
28、[Wanderer](#) (开源)

该项目是用于记录和管理用户的户外探险轨迹的 Web 平台, 帮助你保存珍贵的行程数据。它采用 Go+Svelte 开发, 提供上传、保存、查看 (多种视图) 和分享冒险轨迹的功能, 并支持自托管。



29、[Perplexica](#) (开源)

一个开源的 AI 搜索引擎工具, 灵感来源于 Perplexity AI。它结合了 SearxNG 和大语言模型 (LLMs) 等技术, 能够理解你的问题并深入互联网查找答案, 可作为传统搜索引擎的替代品。

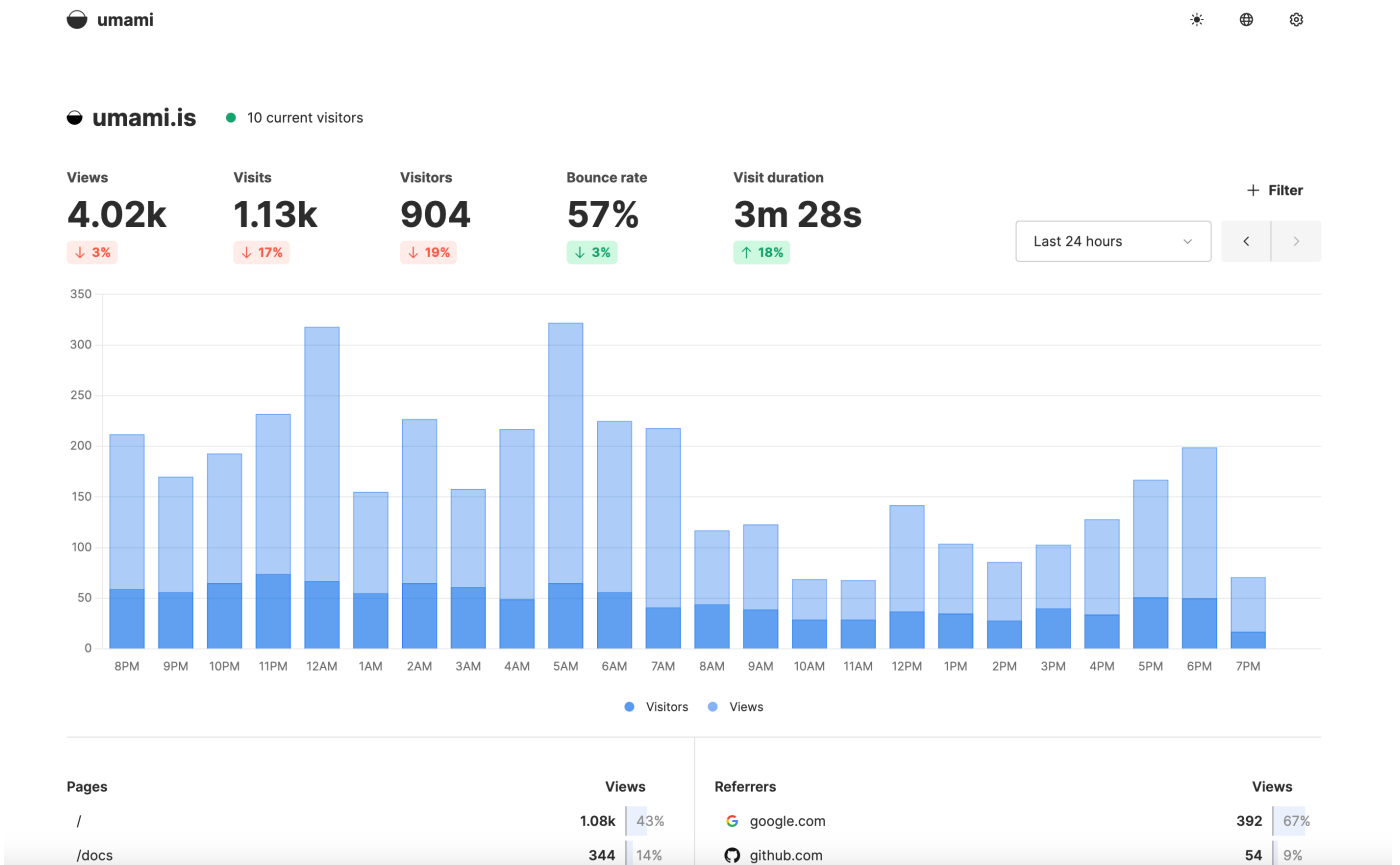


30、[ComfyUI Desktop](#) (开源)

ComfyUI官方提供的桌面客户端，支持Windows、macOS。

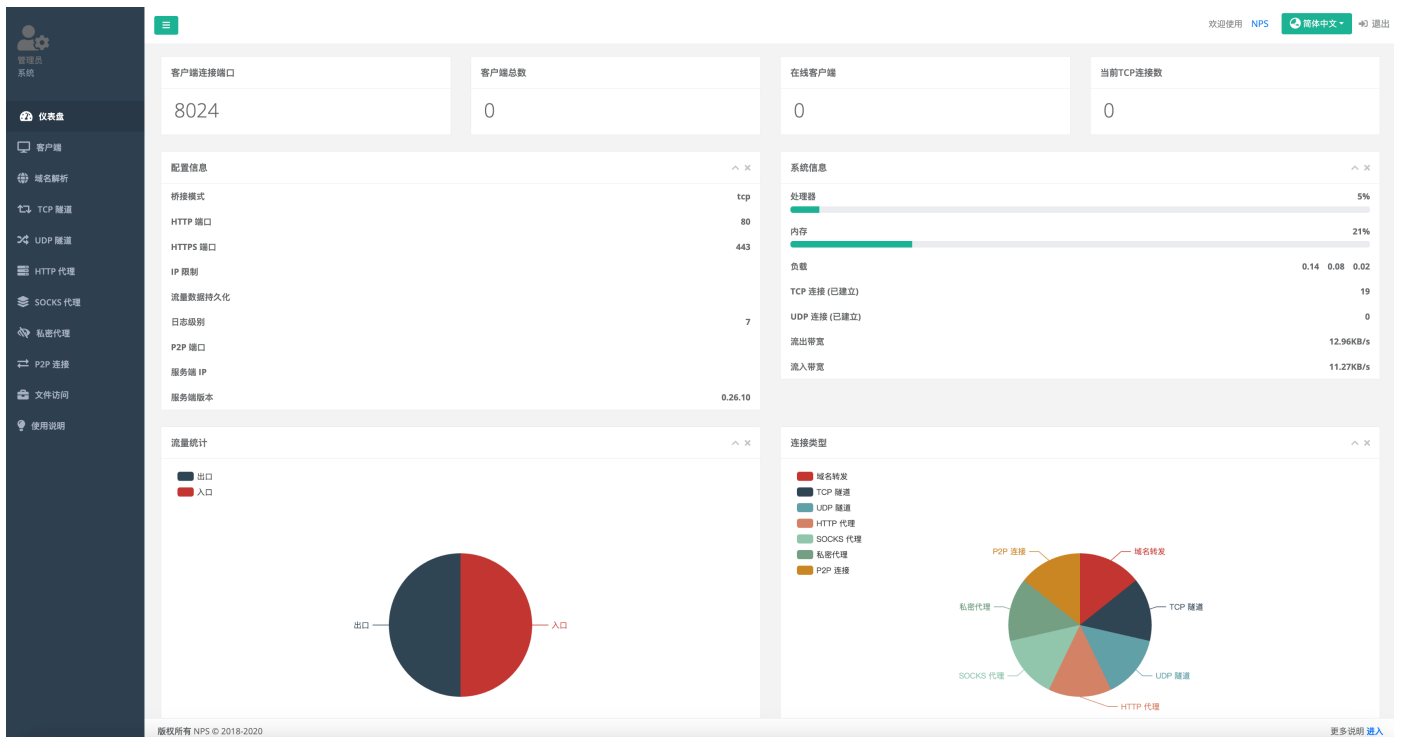
31、[Umami](#) (开源)

Umami是一个简单，易用，可以自行部署的Web站点分析工具。它为用户提供了一个与Google Analytics相比开源、友好、更关注隐私性的站点统计工具，只采集你关心的指标并将全部的内容展示在一个页面上。[官方体验](#)



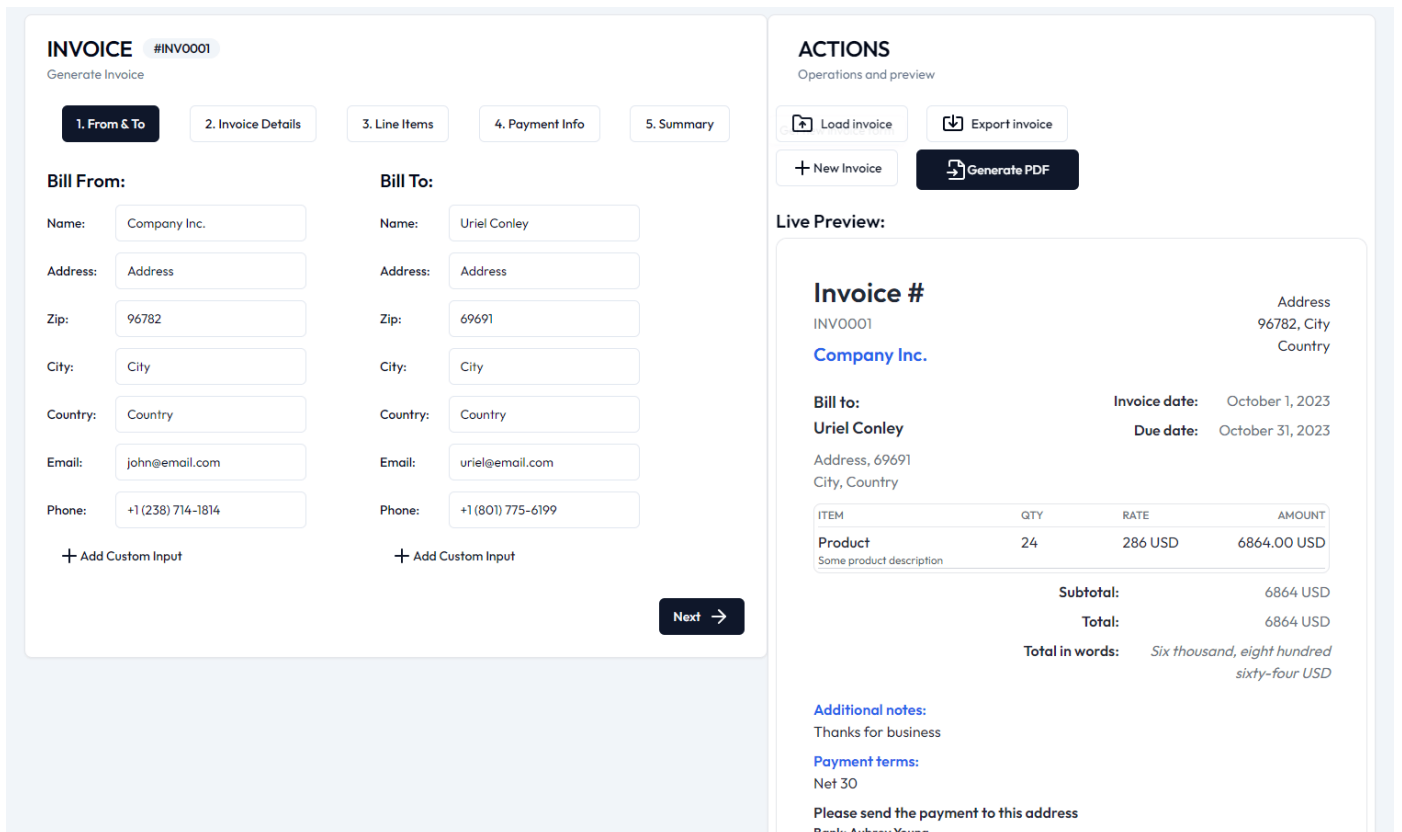
32、NPS (开源)

NPS内网穿透工具是一款轻量级、高性能、功能强大的代理服务器，用于实现内网与外网之间的通信。它可以将外部网络请求转发到内网服务器，使得内网服务器可以被外部访问到。NPS支持TCP和UDP流量转发，并且可以支持任何TCP和UDP上层协议，例如访问内网网站、本地接口调试、SSH访问、远程桌面等。



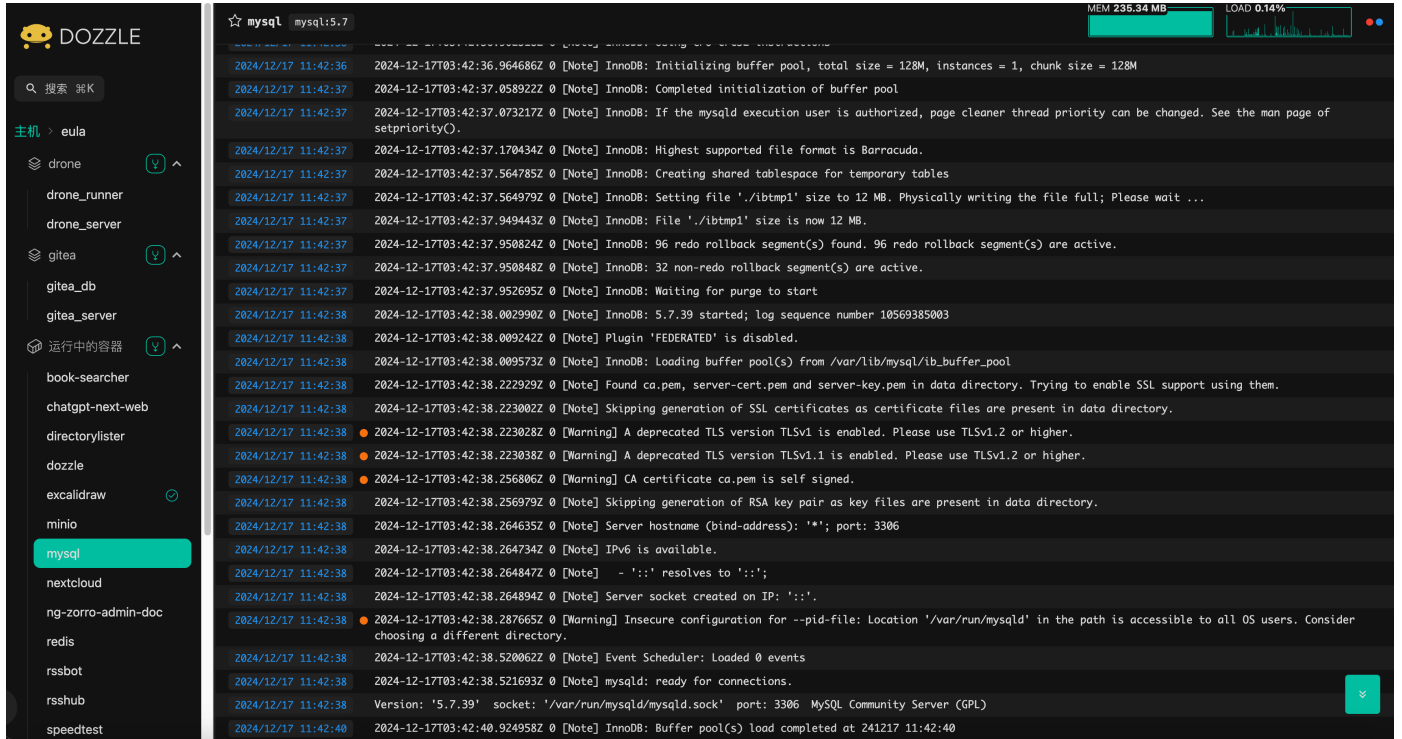
33、Invoify (开源)

invoify 是一个基于 Next.js、TypeScript 和 Shadcn 的在线发票生成应用。



34、Dozzle (开源)

Dozzle 是一款小型轻量级应用程序，具有基于 Web 的界面来监控 Docker 日志。它不存储任何日志文件，仅用于实时监控容器日志。

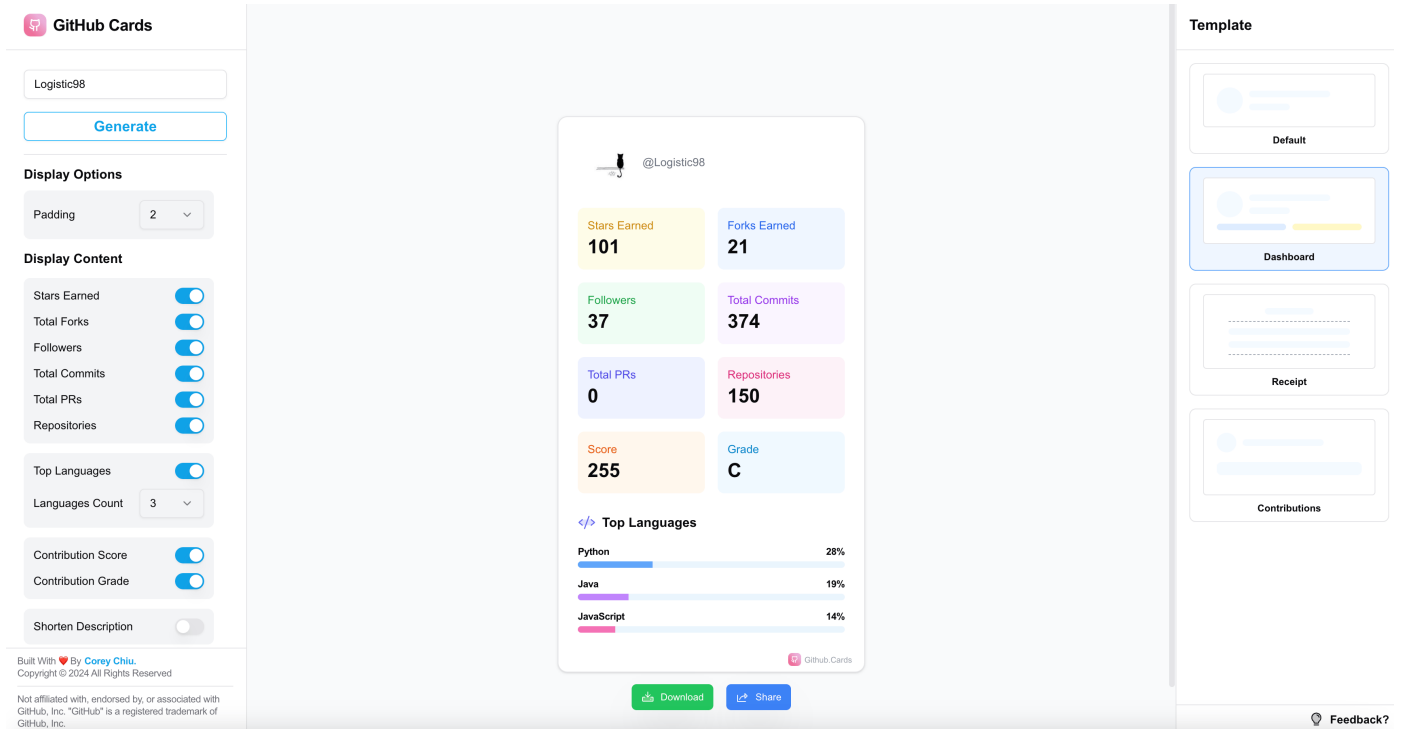


35、[Buzee](#) (开源)

一个桌面软件，提供本地文件的全文搜索（包括图片里面的文字），支持 Windows 和 Mac。

36、[GitHub Cards](#) (免费)

将 GitHub 用户个人数据转换成卡片图。

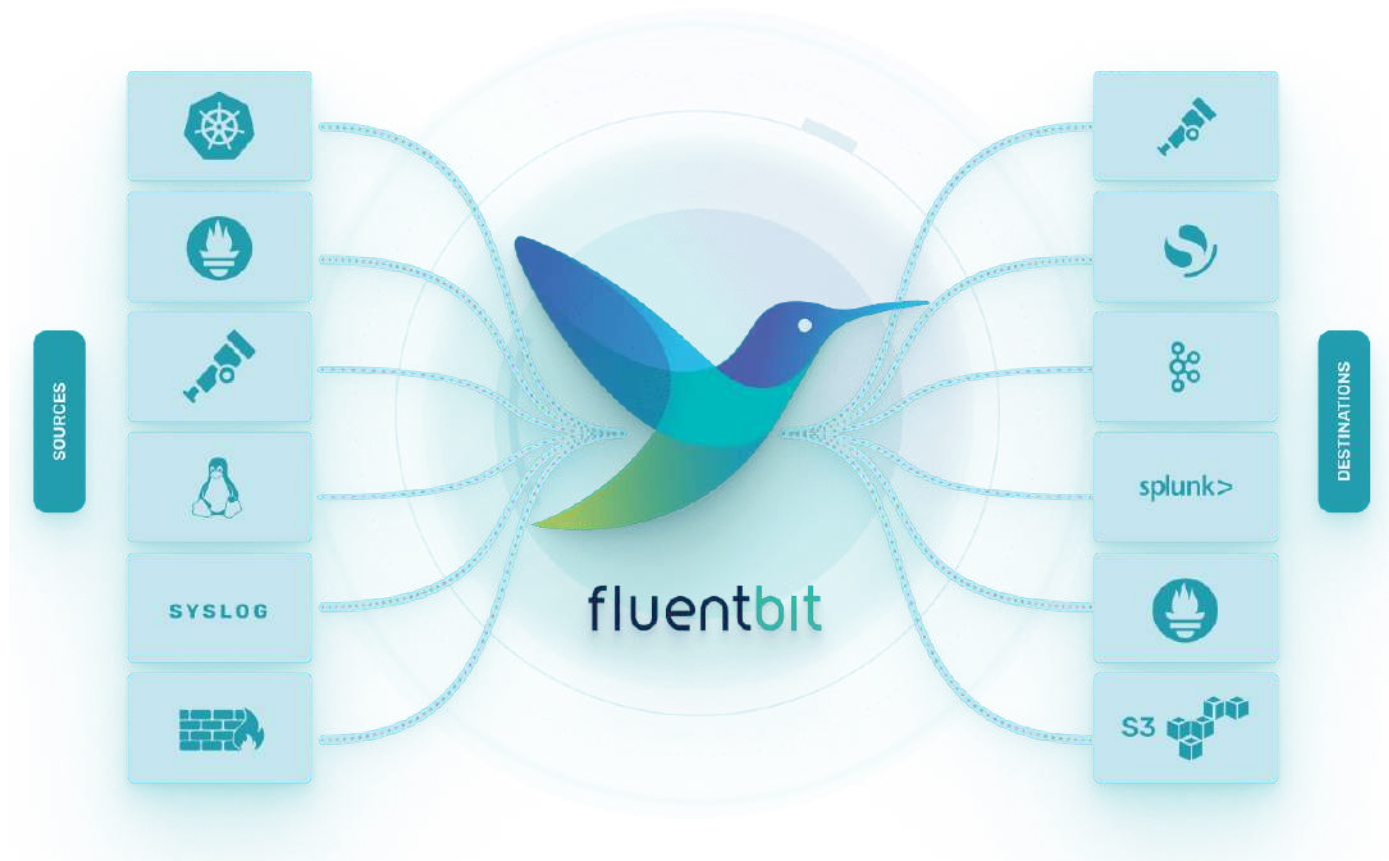


37、[lima](#) (开源)

一款在 macOS 上创建和管理 Linux 虚拟机的工具，它通过 QEMU 和 macOS 原生虚拟化技术，提供一个开箱即用、轻量级的 Linux 虚拟机环境，支持运行各种 Linux 应用和工具，以及 Docker、Podman、K8s 等容器。

38、[fluent-bit](#) (开源)

一个轻量级、高效的日志收集、处理和转发工具，支持实时收集和处理来自不同来源的日志数据，具有低内存和 CPU 占用的特性。



39、[ha_xiaomi_home](#) (开源)

该项目是由小米开源的 Home Assistant 集成插件，可实现小米 IoT 智能设备与 Home Assistant 的无缝集成，支持设备发现、状态同步、远程控制等功能。

40、[docker-xiaoya](#) (开源)

一键部署 Alist、Emby 和 Jellyfin 服务的解决方案，帮你轻松构建完整的家庭多媒体中心，支持 Linux、macOS、Windows 等平台。

41、[jitsi-meet](#) (开源)

这是一个基于 WebRTC 构建的视频会议平台，提供高清音视频质量，支持私聊、举手、投票和虚拟背景等多种功能，兼容所有主流浏览器和移动平台。



42、[runlike](#) (开源)

这是一个用于解析运行中容器的工具，可自动生成对应的 `docker run` 启动命令。它能够提取容器的配置信息，包括包括端口绑定、映射卷、环境变量、网络设置等，适用于复制、调试或迁移容器的场景。

43、[widdershins](#) (开源)

Widdershins用于将符合 OpenAPI 3.0、Swagger 2.0、AsyncAPI 1.x、Semoasa 规范的在线接口文档导出成本地 Markdown 离线文档。

六、学习资源

1、[Prompt_Engineering](#) (英文)

这份教程致力于帮助用户掌握与 LLM 沟通的技巧。内容涵盖从基础到高级的提示工程技术，附有详细的实现指南和示例代码。

2、[开源许可证选择器](#) (中文)

回答几个问题，帮你选择一个开源许可证。

开源许可证选择器

回答这些问题，找到你的真爱：

- 你希望别人用你的代码赚钱吗？
 - 当然，让他们赚钱，我继续吃土！
 - 不行，我的代码是用来改变世界的，不是用来让别人发财的！
 - 呃...我还在纠结要不要当活雷锋
- 你希望别人能修改你的代码吗？
 - 来呀，互相伤害啊！
 - 不行，我的代码完美无缺
 - 我还在纠结中
- 你担心专利流氓吗？
 - 怕死了，我晚上做噩梦都是他们
 - 不怕，我用键盘跟他们决斗
 - 专利流氓是什么东西？能吃吗？
- 你的软件是否会通过网络提供服务？
 - 是的，我要统治互联网！
 - 不，我的软件很害羞，不喜欢出门
 - 可能吧，看心情
- 你在意与其他开源项目的兼容性吗？
 - 非常在意，我想和所有人做朋友
 - 不在意，我的项目就是宇宙
 - 看情况，朋友可以有，但不能太多

清除所有选择

为你推荐的许可证：(0/15)

MIT License

简单宽松的许可证，只要求保留版权和许可声明。

趣味解释
MIT就像自助餐 - 随便吃,但别忘了说是在哪儿吃的!

优点:

- 简单易用
- 高度兼容
- 适合商业用途
- 允许闭源修改
- 无专利条款

缺点:

- 对作者保护有限
- 可能被闭源软件使用
- 无明确专利授权

使用量: 1,400,000 个网站
示例项目: React, Angular
了解更多

Apache License 2.0

平衡了开源和专利保护的许可证。

趣味解释
Apache 2.0就像一个大家都带菜的聚餐,大家分享菜谱,但没人能说自己发明了'食物'这个概念。

优点:

- 专利授权条款
- 商业友好
- 允许闭源修改

缺点:

- 复杂度较高
- 可能与 GPL 不兼容

使用量: 700,000 个网站
示例项目: Android, Kubernetes
了解更多

3、[逆向工程教程](#) (英文)

面向初学者的逆向工程英文电子书，涵盖了 x86 和 ARM 架构。

4、[Web 应用认证实现指南](#) (英文)

这是一本介绍如何在 Web 应用中实现认证的书籍，内容涵盖设计认证流程、存储用户凭据、保护用户数据等方面的指导与建议。

5、[编程十年的感悟](#) (中文)

无论是编程，还是其他的技能，都是「马太效应」，你学得越多，你懂得越多，再学新的东西，你就会学得越快。代码写多了才意识到，程序员的竞争力并不是写代码，也并不是哪门语言或者框架，其核心竞争力是通过技术解决问题的能力，又何必再去拘泥于哪门具体的编程语言或技术呢。

七、生活经验

1、[中国天眼阵列](#)

“中国天眼”是世界最大的射电望远镜之一，直径500米。为了提高分辨率，天文台决定在周边的山头上，再建设24台直径40米的射电望远镜。这样的话，一共25台射电望远镜，组成一个阵列，可以大大提高太空图像的清晰度。

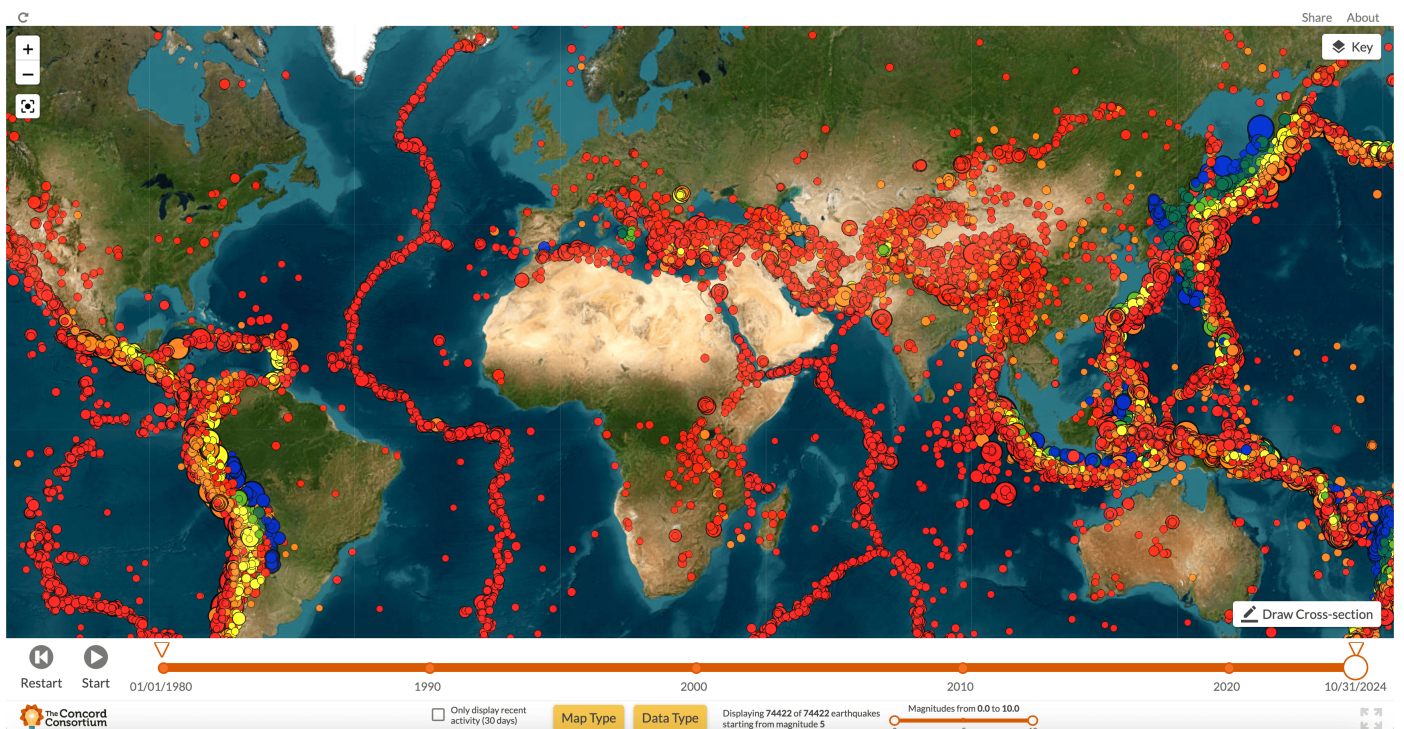


2、[拉格朗日标准模型](#)

拉格朗日标准模型是一个描述粒子行为的统一方程，非常复杂。这个方程有几十行，分成5个部分。尽管如此，它依然是描述粒子行为最简单、最紧凑的形式。

3、[地震地图](#)

这个网站是地震地图，标识出了1980年后大于5.2级的地震，可以清晰地看到地震带，也可以查看某个地区发生过的地震。



4、[深湖冷却](#)

加拿大最大城市多伦多，就在安大略湖旁边。安大略湖是一个很深的湖，最深处达到244米。多伦多市正在建设一个非常酷的项目，使用安大略湖底的冷水，为城市提供冷却系统。安大略湖底部的水温是很低的，该项目从水底85米处取水，即使盛夏，水温也只有4摄氏度。这些水经过40公里的管道，进入市中心的热交换站，为城市带来天然的冷气。



5、[无豆咖啡](#)

一些西方公司最近推出“无豆咖啡”，就是不使用咖啡豆制作的咖啡。它采用废弃的枣子或果核，提取生物质，然后添加提取自茶叶的咖啡因，使其具有咖啡的味道。因此，它比较环保，可以减少食物垃圾，也不必破坏森林种植咖啡树。据说，它的味道非常接近优质咖啡，普通人喝不出差异。只是成本不便宜，略高于传统的优质咖啡豆。



八、闲情逸趣

[为什么说Willow是里程碑式的突破](#)

提到量子计算，美国理论物理学家理查德·费曼曾说过一句令人难忘的话：“大自然不是经典的，如果你想模拟大自然，你最好把它变成量子力学”。在过去30年里，量子计算存在一个根本性的挑战：随着量子比特数量的增加，错误率会急剧上升。然而，这个看似不可逾越的鸿沟，现在终于出现了突破性曙光。

2019年11月23日，谷歌曾在《自然》杂志期刊中发表了一项量子计算突破，谷歌的超导量子芯片Sycamore仅用200秒就完成了一项计算，而世界最快的超级计算机需要1万年。当时谷歌CEO桑达尔·皮查伊（Sundar Pichai）将其比作“莱特兄弟12秒的首飞”。2024年12月10日，谷歌在《自然》发表了最新量子芯片Willow的研究成果，再次取得了里程碑式的突破，主要体现在两个方面：

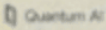
- 首先，Willow实现了随着量子比特数量增加、而指数级降低错误率的目标，通过逐步扩大量子比特阵列规模，从3x3到5x5再到7x7，每次都能将错误率降低一半。这是自1995年Peter Shor引入量子纠错以来，一直是该领域面临的艰巨挑战。
- 其次，更受关注的是其计算能力的突破。在随机电路采样（RCS）基准测试中，Willow用不到5分钟就完成了一项计算，而当今最快超级计算机Frontier则需要 10^{25} 年才能完成。为了让大家直观理解这个数字的概念，谷歌说“这个时间比宇宙的年龄还要大”。


谷歌量子AI创始人兼负责人Hartmut Neven在会上表示：“当我们于2012年创立谷歌量子AI团队时，愿景是构建一个有用的大规模量子计算机，利用我们今天所知的量子力学（自然界的‘操作系统’）来推动科学发现、开发有益的应用、并解决一些社会的关键挑战。”

有意思的是，谷歌量子硬件总监Julian Kelly在简报会上介绍说，此前谷歌量子芯片Sycamore是在加州大学圣巴巴拉分校的一个共享洁净室中建造的——该实验室于2013年宣布成立，为谷歌研究人员提供了更多工具和更强大的功能。而此次，Willow是在谷歌自己的专用超导芯片制造设施生产，可以更好地控制制造工艺参数，良品率和一致性得到提高。Willow基本上继承了Sycamore的所有优点，但实现了更大的里程碑式的突破。


Agenda

- About Google Quantum AI
Hartmut Neven
- Our benchmark performance breakthrough
Hartmut Neven
- Our error correction breakthrough
Michael Newman
- What's notable about Willow
Julian Kelly
- What's next?
Charina Chou
- Q&A







Hartmut Neven
Founder and Lead



Michael Newman
Research Scientist



Julian Kelly
Director of Quantum Hardware



Charina Chou
Director and COO

指数级量子误差校正：低于阈值

量子比特，是量子计算机的运算单位，但它们非常“不稳定”，往往会因周围环境而丢失信息，通常情况下，使用的量子比特越多，错误就会越多——所以，“错误”是量子计算面临的巨大挑战之一。但谷歌这次做到了相反的效果：当谷歌量子芯片Willow中使用的量子比特越多，错误反而大幅减少了，谷歌测试了越来越大的物理量子比特阵列，从3x3编码量子比特网格，扩展到5x5网格，再到7x7网格——每次扩大时，错误率都能减少一半。换句话说，谷歌实现了错误率的指数级降低。

在量子纠错中，涉及将许多物理量子比特放在一起并让它们协同工作，也就是通过创建一个“逻辑量子比特”来纠正错误，3x3、5x5、7x7这样的组合被称为“逻辑量子比特”。1个中心位置的物理量子比特存储实际的量子信息（数据比特），周围8个物理量子比特是辅助比特，这样一个3x3的排列实际上只能存储1个比特的信息，但它能够保护这个信息不被环境干扰破坏。就好比是运输一个易碎品（量子信息），中心是易碎品本身（数据比特），周围8个位置是包装泡沫（辅助比特），虽然看起来用了9个空间位置，但实际运输的有效物品只有中心的那一个，但这些都“包装泡沫”让运输变得更安全可靠。

这就解释了为什么量子计算机需要这么多“物理量子比特”，而且物理量子比特数量看起来很多，但实际能用于计算的“逻辑量子比特”数量要少得多：比如要存储10个比特的信息，使用3x3的方案就需要90个物理量子比特(10x9)，用5x5方案则需要250个物理量子比特(10x25)，用7x7方案需要490个物理量子比特(10x49)。这种“冗余”是必要的，因为它保证了量子计算的可靠性。

“我们希望随着这些集合越来越大，纠错能力也越来越强，这样量子比特就会越来越准确。问题是，随着这些东西越来越大，出错的机会也越来越多，所以我们需要足够好的设备，这样当我们把这些东西做得越来越大时，纠错能力就能克服我们引入系统的这些额外错误。”谷歌实验室研究科学家Michael Newman在简报会上说。谷歌称，这是一个30年来一直未实现的目标，直到现在Willow实现了突破——实现了每次逻辑量子比特的大小增加，从3x3到5x5再到7x7，错误率就会呈指数下降。这就像是在搭积木，以前积木堆得越高就越容易倒，但现在谷歌的这项研究，不但让积木能堆得更高，而且越高反而越稳固。这也就有力地表明，未来实用的超大型量子计算机确实可以构建。

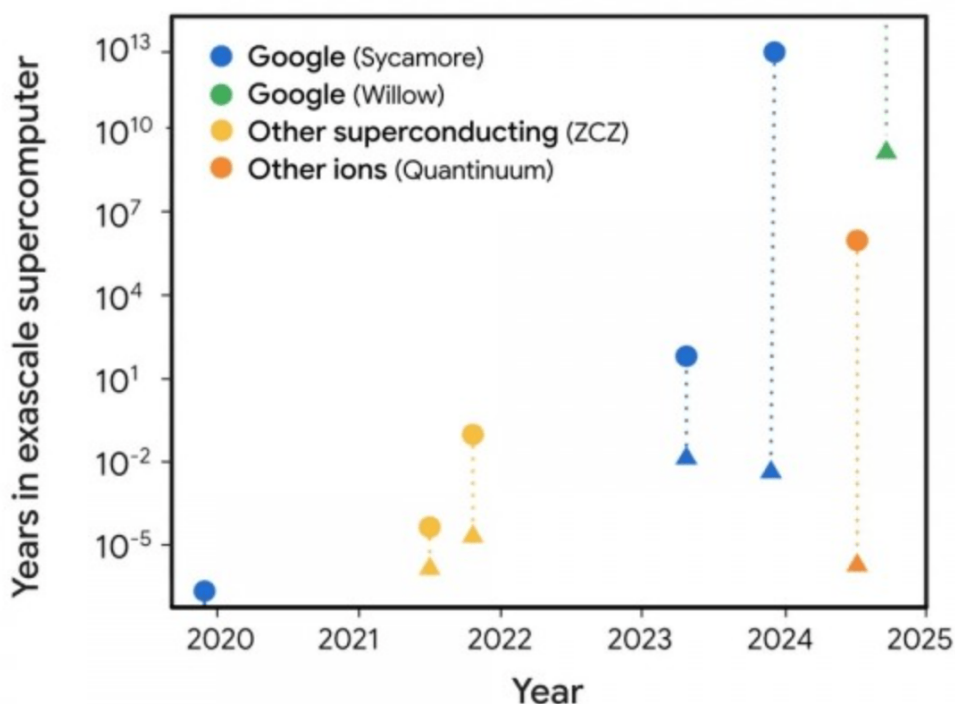
这一突破在业界被称为“低于阈值”——即能够在增加量子比特数量的同时降低错误。在《自然》杂志的这篇论文中，研究人员写道：“虽然许多平台已经展示了量子纠错的不同特性，但至今没有一个量子处理器明确地表现出低于阈值的性能，如果不低于阈值，那么进行量子纠错真的毫无意义，这确实是未来实现这项技术的关键因素”。Julian Kelly补充说道：“量子比特本身的质量必须足够好，才能进行纠错，我们的纠错演示表明，在集成系统层面，一切都同时工作，这不仅仅是量子比特的数量、T1 或双量子比特错误率的问题。这也是这项挑战长期以来一直难以解决的原因之一”。Hartmut Neven说道：“Willow让我们更接近运行传统计算机上无法复制的实用、商业相关算法”。

5分钟完成一次计算，而Frontier则需要 10^{25} 年

为了衡量 Willow 的性能，谷歌使用了随机电路采样 (RCS, Random Circuit Sampling) 基准。Hartmut Neven介绍说：“RCS 由谷歌量子AI团队首创，现已被广泛用作该领域的标准，是当今量子计算机领域最难的经典基准”。

具体而言，RCS被用来展示量子计算机和经典计算机之间快速增长的差距，并强调量子处理器如何以双指数速度剥离，并将随着量子位的扩大而优于经典计算机。它涉及产生和测量随机量子电路的输出（随机量子电路是以一种看似任意的方式应用于量子位的量子门序列）。

如开篇所述，Willow在RCS测试中的表现令人惊叹：它在不到五分钟的时间内完成了一项计算，而当今最快的超级计算机Frontier需要 10^{25} 年。“它证实了量子计算发生在许多平行宇宙中的观点，这与 David Deutsch 首次提出的‘我们生活在多元宇宙中’观点一致”。



在简报会上被问及“在实际应用下，我们离看到量子计算机还有多远”时，Hartmut Neven介绍说，量子计算机在药物发现、核聚变反应堆、肥料生产、量子机器学习、电动汽车电池等都有用武之地。

- 在药物发现方面。“约75%的小分子药物都会被P450酶代谢，这基本上是小分子药物必须避开的一个关卡，这个酶目前还没有被充分理解，而量子计算机有望能更好地对它建模，谷歌正在研究这一应用，试图用量子计算机理解酶复合物P450。”
- 在机器学习方面。“现在AI无处不在，但重要的是要认识到，有许多基础和计算问题，比如解决困难的优化问题或分解大数 (Integer Factorization)，这些无法只通过学习来解决，因为你需要庞大的训练数据。这也是量子计算机能够帮忙的地方。”

谷歌量子AI主任兼首席运营官Charina Chou补充说道：“现在的AI主要指机器学习，机器学习需要大量训练样本。比如ChatGPT的惊人成功，是因为有大量可用的训练数据。在这方面，量子计算也能帮忙。谷歌实际上在这方面已经有一些展开工作，这将给我们带来能从磁共振成像和核磁共振中获得更多价值的算法。这些新的量子算法可以作为一个原子尺子，给出分子中原子核之间非常精确的距离。所以量子计算可以帮助收集原本无法获取的训练数据集，这是其与AI的另一个重要联系”。

此外，Charina Chou还指出，“模拟大自然的最大机会可能就在量子力学系统中”，谷歌正在与许多大公司、学术机构和初创公司在物理、化学、材料科学领域展开合作，探索量子计算在各领域的应用场景。



系统工程是关键

在Hartmut Neven看来，系统工程是设计和制造量子芯片的关键：芯片的所有组件，例如单量子比特门和双量子比特门、量子比特复位和读出，都必须同时经过精心设计和集成。如果任何组件滞后或两个组件不能很好地协同工作，就会拖累系统性能。“因此，最大化系统性能贯穿于我们流程的各个方面，从芯片架构和制造到门开发和校准。Willow取得的成果是整体评估量子计算系统，而不是一次只评估一个因素。”

目前，Willow在上述两个系统基准测试中（量子纠错和随机电路采样）均拥有一流的性能，除此之外，Willow的T1时间（测量量子比特可以保留激发的时间长短——关键的量子计算资源）接近100 μ s，比Sycamore芯片的20微秒提了5倍。

当被问到“从2019年53量子比特的Sycamore，到现在105量子比特Willow的新成果，谷歌在量子计算上的技术路线似乎更注重质量而非数量，这是否意味着业界普遍追求‘更多量子比特’的路线需要调整？”，这个问题时，Hartmut Neven是这样回复的：

- 量子计算机需要同时具备“数量”和“质量”两个条件。简单地增加量子比特数量是不够的，因为如果错误率太高，这些量子比特就无法被有效利用。这就像是一台电脑，如果经常死机，即使配置再高也无法正常工作。
- 如果一个量子计算机的门操作错误率是千分之一，那么执行一千次操作后，系统就很可能出错。而在实际应用中，每个量子比特至少需要执行十次门操作。所以对于一个有100个量子比特的系统来说，需要将错误率控制在十万分之一才算合格。
- 相比之下，某些其他设计虽然号称有上千个量子比特，但错误率高达1/200。这种情况下，在整个系统崩溃之前，根本无法同时使用所有量子比特。“这就是为什么谷歌选择先把重心放在提高量子比特的‘质量’上，因为只有先解决了质量问题，增加数量才有意义”。

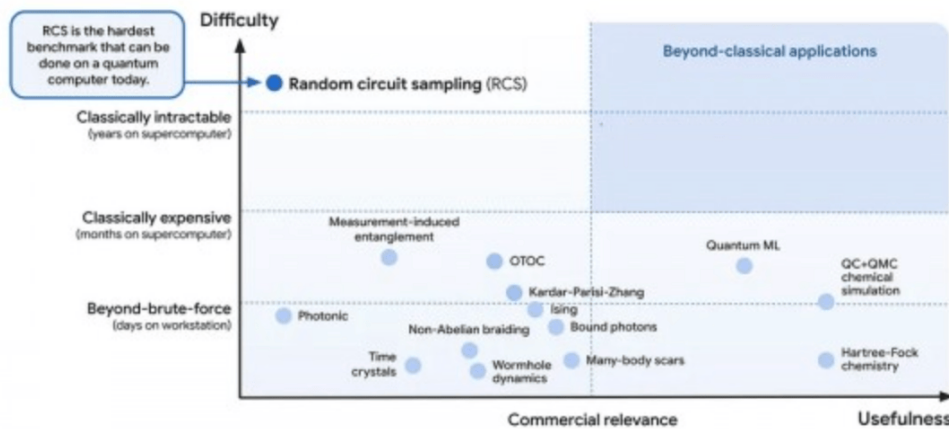
谷歌的研究团队表示，他们正在开发新的技术来扩大系统规模。当前的重点是降低错误率，让它达到量子纠错的要求。随着技术的成熟，量子比特的数量也会逐步增加。

谷歌的量子计算之旅

到目前为止，谷歌针对量子计算进行了两种不同类型的实验。一方面，运行RCS基准测试，该基准测试衡量了与传统计算机的性能，但没有已知的实际应用。另一方面，对量子系统进行了科学模拟，这也带来了一些新的科学发现，但这些发现仍在传统计算机的范围内。

Random circuit sampling (RCS): in context

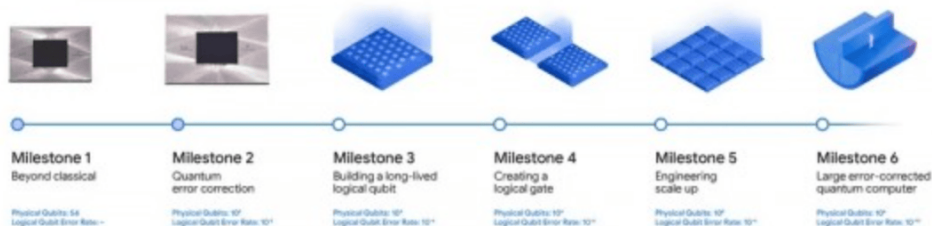
To date, no quantum computer has outperformed a supercomputer on a commercially relevant application. Our latest research is a step towards that direction.



视频简报会上，谷歌量子计算AI团队公布了谷歌量子计算路线图，谷歌称，该路线图的重点是通过开发能够进行复杂、纠错计算的大型计算机，来释放量子计算的全部潜力，这些里程碑将引领我们朝着有意义应用的高质量量子计算硬件和软件发展。图上可见，该路线图包含六个里程碑，谷歌目前的已经完成了两个里程碑节点。

Our quantum computing roadmap

Our focus is to unlock the full potential of quantum computing by developing a large-scale computer capable of complex, error-corrected computations. We're guided by a roadmap featuring six milestones that will lead us toward top-quality quantum computing hardware and software for meaningful applications.



谈及投身于这趟量子计算之旅，Hartmut Neven在谷歌官网上写道：

- “我的同事有时会问我，为什么离开蓬勃发展的人工智能领域，转而专注于量子计算。我的回答是，这两项技术都将被证明是我们这个时代最具变革性的技术，但先进的人工智能将从量子计算中受益匪浅。这就是我将我们的实验室命名为量子人工智能的原因”。
- “量子算法具有基本的缩放定律，就像我们在RCS中看到的那样，许多对人工智能至关重要的基础计算任务也具有类似的缩放优势。因此，量子计算对于收集传统机器无法访问的训练数据、训练和优化某些学习架构、以及对量子效应很重要的系统进行建模将是必不可少的。这包括帮助我们发现新药、为电动汽车设计更高效的电

池，以及加速核聚变和新能源替代品的进展。许多未来改变游戏规则的应用程序中在传统计算机上是行不通的，它们正等待量子计算来解锁”。

九、数字与言论

1、工程团队可以分为防守和进攻两种角色。进攻角色负责开发新产品和增加功能，防守角色负责维护产品。这两个角色都很重要，但是公司往往更重视进攻角色。——[《将工程团队分为防守和进攻》](#)

2、简单的东西很难卖出去，因为简单的东西看起来很容易，谁愿意为容易买单呢？商人们早已明白这个秘密，容易的东西往往被包装成复杂的东西，让用户觉得它很特别，从而为这种特别性付出高价。——[《贩卖复杂性》](#)

3、创业者的工作重点始终是优化业务，而不是优化代码库，除非代码库就是你的业务。——[《使用 HTMX 构建 SaaS 产品》](#)

4、我热爱消费者市场，讨厌企业市场。我们推出了一种产品，告诉每个人，大家自己决定要不要买，这很简单。但是，企业市场不是这样，使用产品的人自己做不了主，而做主的人不使用产品。——[乔布斯](#)

5、一项美国的调查发现，81%的公司发布过虚假的招聘广告，目的有几个：（1）在招聘网站持续亮相，（2）建立人才库，（3）让现有员工感到自己是可替代的，从而不敢松懈。——[《幽灵工作对科技从业者的破坏》](#)

6、现代硬件的速度快得难以置信，苹果的 M1 Max 芯片运行频率为 3.2GHz，即每秒32亿个时钟周期。然而，我用它打开微软的 Teams 需要3秒，我不相信这需要花费96亿个时钟周期。——[preyneyv.dev](#)

7、美国心理学会的一项研究表明，消耗脑力的思考，会导致烦躁、沮丧或其他负面情绪，所以“思考不利于健康”这个观点是对的。——[合众社报道](#)

8、AI 研究已经70年了，人们最近才发现，让 AI 成功的真正方法是大量的计算。根本原因就是摩尔定律，它使得计算成本持续呈指数式下降。——[《马斯克的梦想和 AI 的教训》](#)

9、我遇到一个程序员，他完全不想使用 AI。因为他用过一次，发现错误很多，就不想再用了，也不想了解。这件事令我很困惑。我理解 AI 被过度炒作了，可能对你正在做的事情没有用处。我不理解的是，在2024年，你怎么可能成为一名程序员，却对一项据说会从根本上改变我们未来编程方式的技术一点也不好奇。——[《他们都使用它》](#)

10、凡是让你顾全大局的，你一定不在这个局内。凡是不惜一切代价的，你往往就是这个代价。凡是禁止的，往往是有好处的，但是不能和你分享。凡是提倡的，往往是有大坑的，且需要你去填大坑。只要你能吃苦，就有吃不完的苦。只要你能吃亏，就有吃不完的亏。——[Linux.Do](#)

11、产品经理的职责就是三件事：（1）了解用户需求；（2）提出解决方案；（3）安排任务执行。——[《产品经理的 AI》](#)

12、伟大的科学家是想法大胆的人，但是对自己的想法同时持有高度批判的态度。接受自己的猜想之前，他们首先会严厉尝试反驳它。只有发现很难反驳成功时，他们才相信自己的想法是对的。——[卡尔·波普尔](#)，科学哲学家