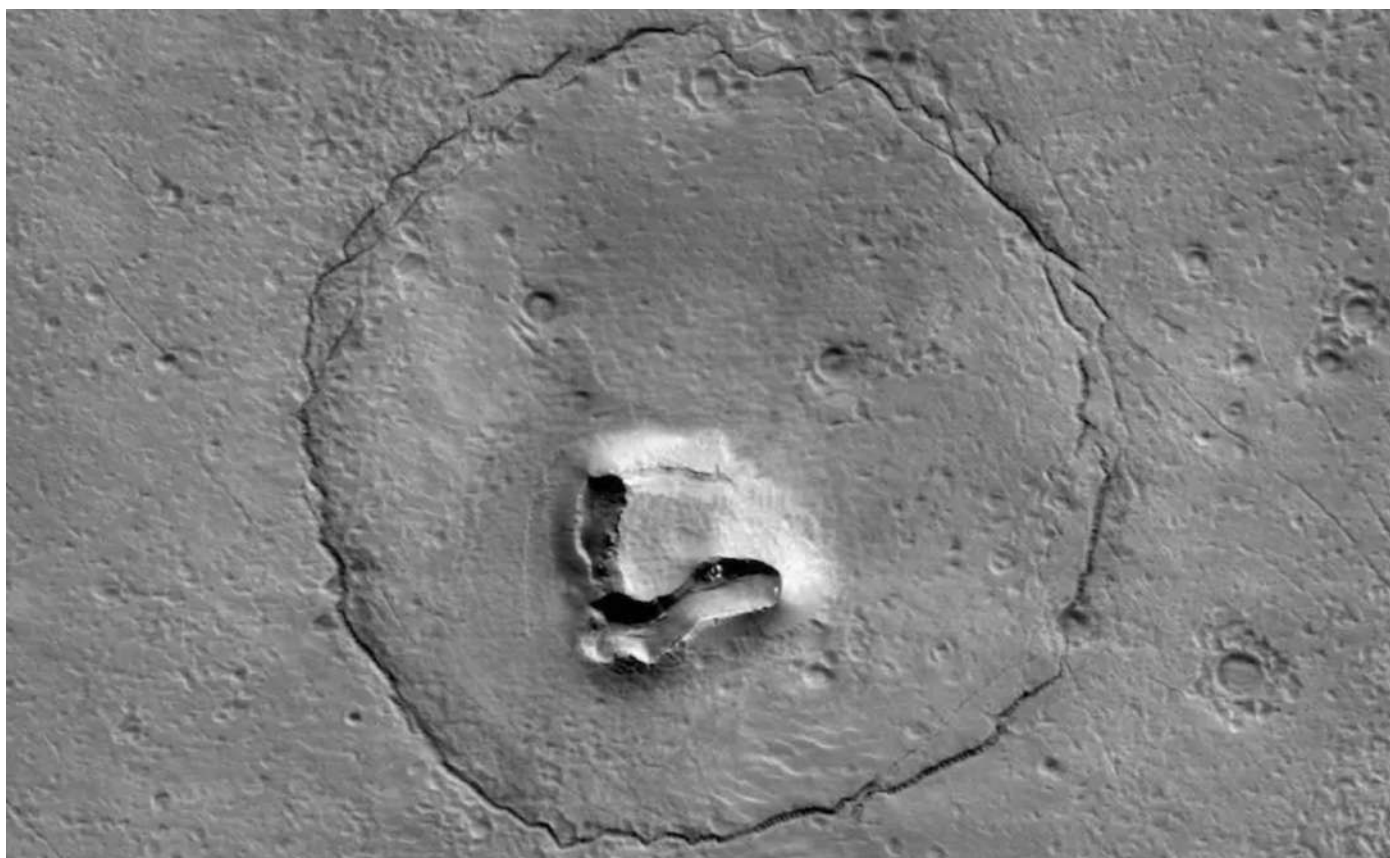


# 2023.01-03--第11期

---

## 一、刊首图

---



美国宇航局公布了一张 [火星地表](#) 照片，看上去有点像泰迪熊，熊脑袋的直径大约2000米。

## 二、时事新闻

---

### 1、[暴雪国服停服](#)

2023年1月24日0时，由网易代理的《魔兽世界》、《炉石传说》、《守望先锋》、《暗黑破坏神III》、《魔兽争霸III：重制版》、《风暴英雄》、《星际争霸》系列产品在中国大陆市场的所有运营将正式终止。届时，暴雪将关闭战网登录以及所有游戏服务器，同时关闭客户端下载，将于停服后公布暴雪游戏产品的退款工作安排。

### 2、[最后一架波音747](#)

波音公司交付了最后一架波音 747 飞机，从此不再生产该型号。作为纪念，这架飞机在天空中画出了“747”的飞行轨迹，地面通过接收飞机的 ADS-B 位置信号，就可以看到。波音 747 是世界上第一架双通道飞机，于1968年首次制造，美国总统座机“空军一号”就是改装的波音 747。由于它成本高、油耗大，波音计划在2025年用更省油的双引擎波音 777X 取代它。



### 3、[禁止短途航班](#)

法国通过一项法律，为了环保，规定如果两个城市之间有直通的火车，并且火车行程不多于2.5个小时，那么将禁止这两个城市之间的飞机航班。这是为了鼓励多采用火车，少使用耗能大的飞机。法国的很多大城市都会因此受影响，比如巴黎到波尔多、巴黎到里昂、里昂到马赛，航班可能都会取消。如果以高铁每小时300公里计算，2.5小时的行程就相当于750公里，等于北京到徐州的距离，只有超出这个距离，才可以坐飞机。

### 4、[Meta开源LLaMA](#)

Meta 首席 AI 科学家 Yann LeCun 宣布，他们「开源」了一个新的大模型系列——[LLaMA](#) (Large Language Model Meta AI)，参数量从 70 亿到 650 亿不等。这些模型的性能非常优异：具有 130 亿参数的 LLaMA 模型「在大多数基准上」可以胜过 GPT-3（参数量达 1750 亿），而且可以在单块 V100 GPU 上运行；而最大的 650 亿参数的 LLaMA 模型可以媲美谷歌的 Chinchilla-70B 和 PaLM-540B。

### 5、[OpenAI 发布 GPT-4](#)

OpenAI发布了GPT-4，它是一个超大的多模态模型，它的输入可以是文字（上限2.5万字），还可以是图像，现在已经可以理解论文中的图表了，在各种专业和学术基准上和人类相当。比如模拟律师考试，GPT-4取得了前10%的好成绩，相比之下GPT-3.5是倒数10%；做美国高考满分为800分的SAT试题，GPT-4也在阅读写作中拿下710分、数学700分的高分。

### 6、[GPT-4 Copilot X发布](#)


GitHub官宣：基于GPT-4的新一代代码生成工具Copilot X来了。新一代Copilot X，直接整合里一个华丽的聊天界面，不仅动动嘴皮子就能生成代码，还能边写代码边跟AI对话，以及为你量身定制的智能文档系统等等。

具体来说，Copilot X提供支持的体验有：Copilot Chat；Copilot for Pull Request；Copilot for Docs；Copilot for CLI。这些新功能都是由OpenAI的GPT-4驱动的。值得注意的是，由于速度延迟的原因，代码自动补全工具仍基于GitHub的Codex模型上，该模型是在GPT-3上训练的。现在，微软已经将GPT-4集成到搜索、办公、写代码等各种真正意义上的生产力工具上，属实是拥有了开启第四次科技革命的力量。



GitHub   
@github

GitHub Copilot is already helping developers code faster in their IDEs. But what's next?

Our answer is GitHub Copilot X. It's our vision for the future of AI-powered software development. Check it out 



github.blog

GitHub Copilot X: The AI-powered developer experience | The GitHub Blog

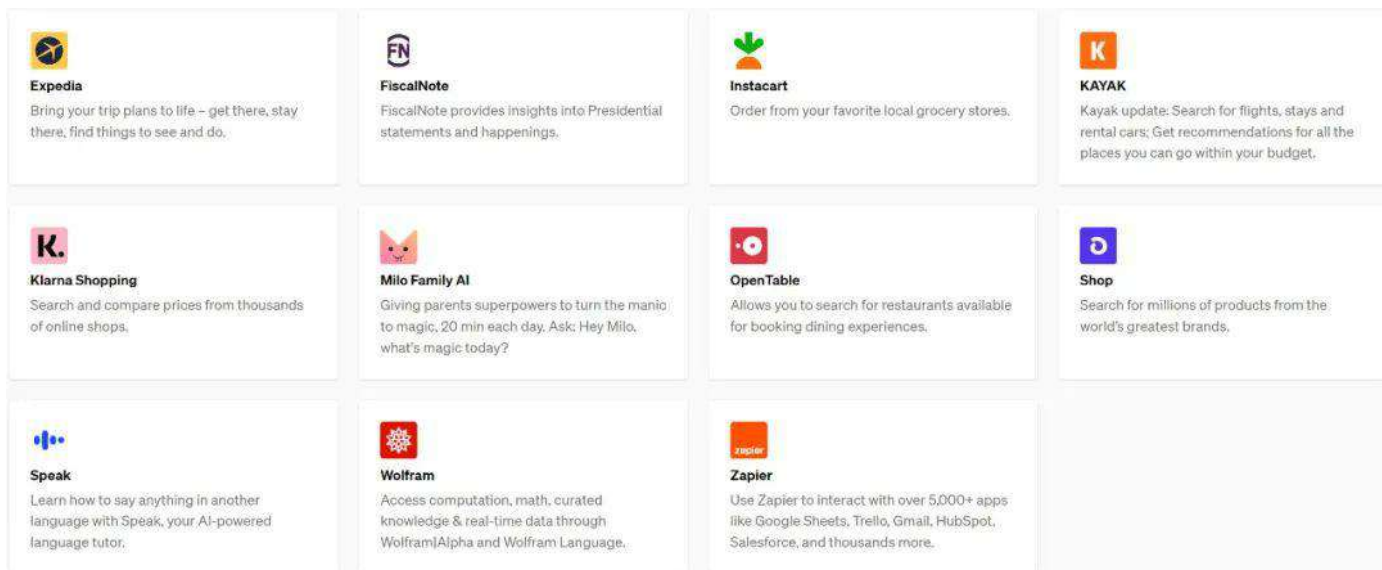
GitHub Copilot is evolving to bring chat and voice interfaces, support pull requests, answer questions, and adopt OpenAI's GPT-4.

10:00 pm · 22 Mar 2023 · **1.3M** Views

## 7、[ChatGPT开始联网](#)

ChatGPT 是个「智商超高」的人工智能，经过 GPT-4 内核的升级版更是如此。然而我们知道，它的训练数据截止到 2021 年，没有训练过的东西，AI 是不可能准确回答出来的。自去年 11 月上线以来，ChatGPT 已被无数人使用，人们一直要求以各种形式让这个大语言模型接入更多数据，3 月 24 日，OpenAI 终于宣布部分解除了 ChatGPT 无法联网的限制。

OpenAI 的解决方案是用第三方插件作为桥梁，让 AI 在较安全的环境下「看」外界数据。昨天该机构开放了第一批 ChatGPT 插件名单。这批插件由 Expedia、FiscalNote、Instacart、KAYAK、Klarna、Milo、OpenTable、Shopify、Slack、Speak、Wolfram 和 Zapier 创建。



具体来说，插件现在可以让你用 ChatGPT 执行以下操作：

- 检索实时信息：例如体育比分、股票价格、最新消息等；
- 检索知识库信息：例如公司文件、个人笔记等；
- 代表用户执行操作：例如，订机票、订餐等。

除此之外，OpenAI 还自己提供了两种插件，包括一个网络浏览器和一个代码解释器，并开源了一个知识库检索插件的代码。现在，任何开发人员都可以自行构建插件，用来增强 ChatGPT 的信息库了。

目前插件 alpha 版的访问权限已扩展至更多候补名单中的用户和开发人员，虽然 OpenAI 表示优先考虑少部分开发者和 ChatGPT Plus 用户，但也计划未来进行更大规模的开放。

## 三、技术文章

### 1、[如何修改 Git 仓库的历史](#) (英文)

修改 git 历史是常见需求，但是很难操作。本文介绍使用一个叫做 BFG 的工具，从 Git 仓库删除内容、修改历史。

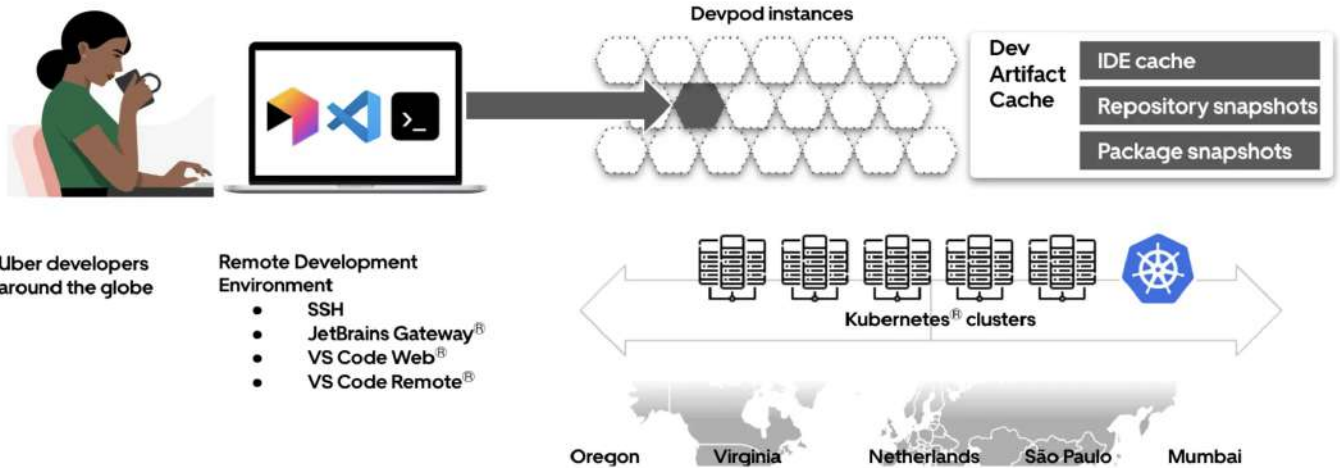
### 2、[Uber的远程开发基础架构](#) (英文)

本文是 Uber 官方文章，介绍他们为了适应全世界的员工远程办公，而搭建的开发基础架构，所有工程师的代码都放在一个单一仓库里面。



# Devpod

## Remote Development Environment @ Uber



### 3、[训练个中文版ChatGPT](#) (中文)

LLaMA 是一系列模型的总称，参数量从 70 亿到 650 亿不等，其中，130 亿参数的 LLaMA 模型「在大多数基准上」可以胜过参数量达 1750 亿的 GPT-3。不过，该模型并没有经过指令微调 (instruct tuning)，因此生成效果较差。为了提高模型性能，来自斯坦福的研究者帮助其完成了指令微调的工作，训练了一个名为 Alpaca 的 70 亿参数新模型 (基于 LLaMA 7B)。

对于普通研究者来说，这是一种切实可行的廉价微调方式，不过需要的运算量仍然较大 (作者表示他们在 8 个 80GB A100 上微调了 3 个小时)。而且，Alpaca 的种子任务都是英语，收集的数据也都是英文，因此训练出来的模型未对中文优化。

为了进一步降低微调成本，另一位来自斯坦福的研究者 ——Eric J. Wang 使用 LoRA (low-rank adaptation) 技术复现了 Alpaca 的结果。具体来说，Eric J. Wang 使用一块 RTX 4090 显卡，只用 5 个小时就训练了一个和 Alpaca 水平相当的模型，将这类模型对算力的需求降到了消费级。而且该模型可以在树莓派上运行。

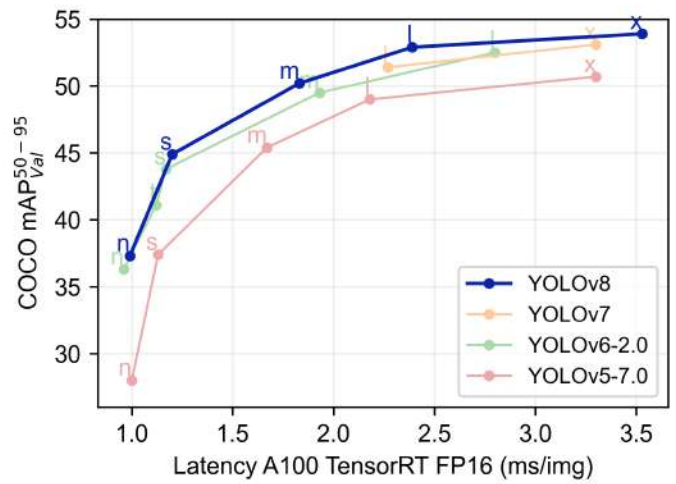
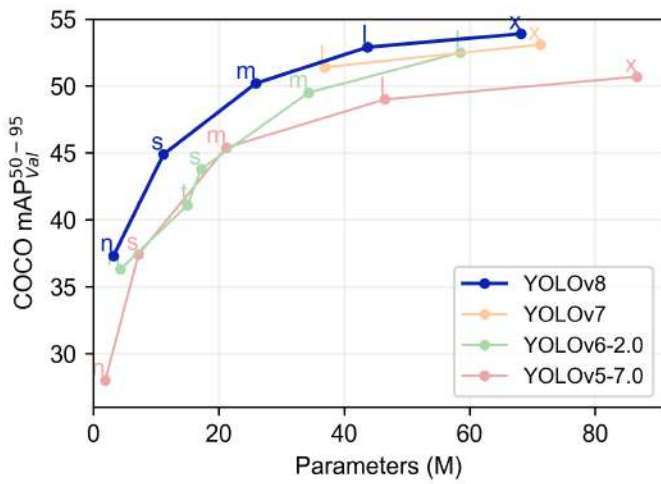
在 [Alpaca-LoRA](#) 项目中，作者提到，为了廉价高效地进行微调，他们使用了 Hugging Face 的 PEFT。PEFT 是一个库 (LoRA 是其支持的技术之一)，可以让你使用各种基于 Transformer 的语言模型并使用 LoRA 对其进行微调。好处在于，它允许你在一般的硬件上廉价而有效地微调模型，并有较小的 (也许是可组合的) 输出。尽管 Alpaca 性能可以媲美 GPT 3.5，但其种子任务都是英语，收集的数据也都是英文，因此训练出来的模型对中文并不友好。

来自华中师范大学等机构的三位个人开发者开源的中文语言模型 [Chinese-alpaca-lora](#)，该项目基于 LLaMA、Stanford Alpaca、Alpaca LoRA、Japanese-Alpaca-LoRA 等完成，单卡就能完成训练部署。这个模型是在 Meta 开源的 LLaMA 基础上，参考 Alpaca 和 Alpaca-LoRA 两个项目，对中文进行了训练。

## 四、开源组件

### 1、[ultralytics](#)

由 Ultralytics 开发的 Ultralytics YOLOv8 是一种尖端的、最先进的 (SOTA) 模型，它建立在以前成功的 YOLO 版本的基础上，并引入了新的功能和改进，以进一步提高性能和灵活性。YOLOv8 旨在快速、准确且易于使用，使其成为广泛的对象检测、图像分割和图像分类任务的绝佳选择。



## 2、[yolov8\\_tracking](#)

使用 YOLOv8 进行实时多目标跟踪和分割。

## 3、[nanoGPT](#)

用于训练/微调中型 GPT 的最简单、最快的存储库。

available GPT implementations

~~minGPT~~ nanoGPT



## 4、[ChatRWKV](#)

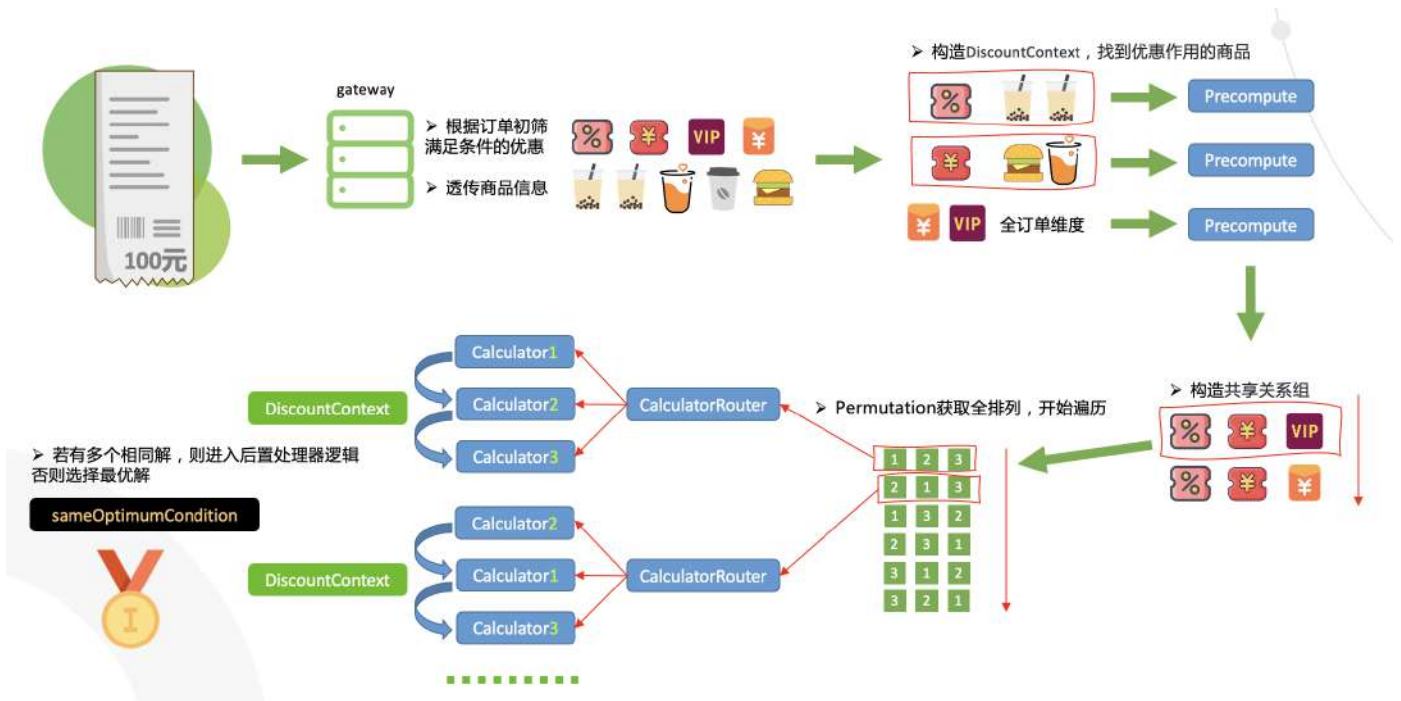
ChatRWKV 类似于 ChatGPT，但由 RWKV（100% RNN）语言模型提供支持，是可以在质量和缩放方面与 Transformer 相媲美的 RNN，同时速度更快并节省 VRAM。

## 5、[plato](#)

腾讯图计算TGraph正式开源高性能图计算框架Plato，达到业界领先水平，超大规模图计算全面进入分钟级时代。

## 6、[Q-calculator](#)

高性能优化计算框架，用来计算复杂场景的优惠折扣。

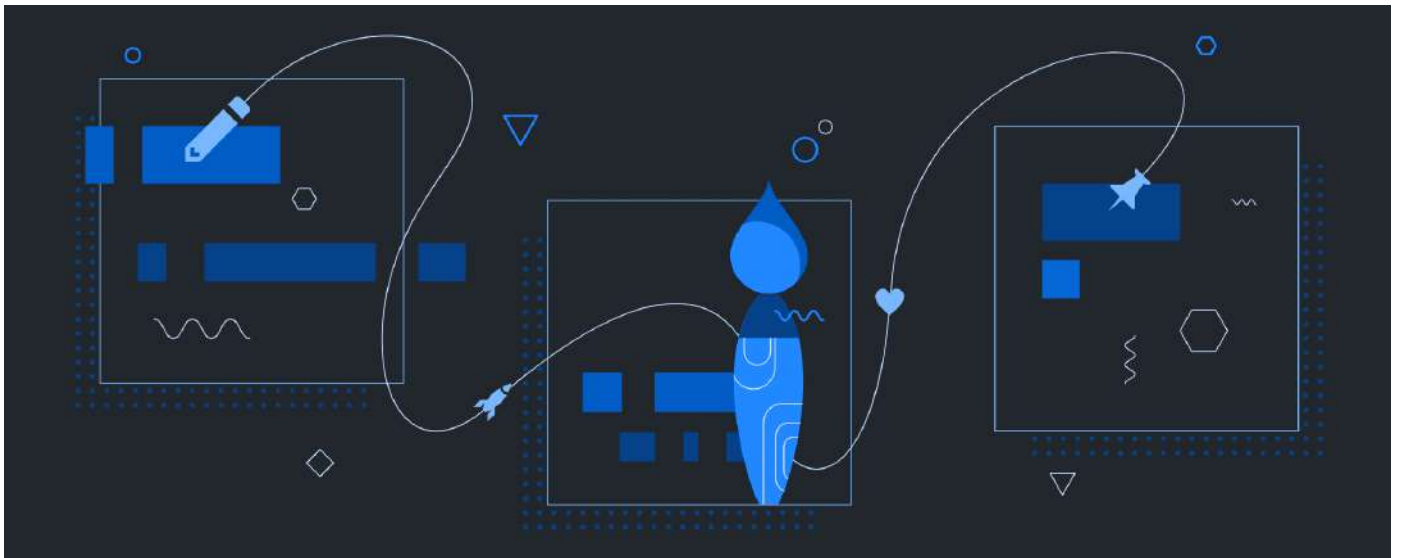


## 7、[point-e](#)

这是 [Point-E: A System for Generating 3D Point Clouds from Complex Prompts](#) 论文的官方代码和模型发布。

## 8、[css](#)

GitHub 开源的设计系统。由 GitHub 设计团队开源和维护的项目，包含了 GitHub 的界面设计原则、使用指南和开箱即用的 UI 组件。



## 9、[QuestPDF](#)

QuestPDF 是用于生成 PDF 文档的现代开源 .NET 库。提供由通俗易懂的 C# Fluent API 提供支持的综合布局引擎。

## 10、[pynecone](#)

Pynecone 是一个全栈 Python 框架，可以在几分钟内轻松构建和部署 Web 应用程序。

## 11、[htslib](#)

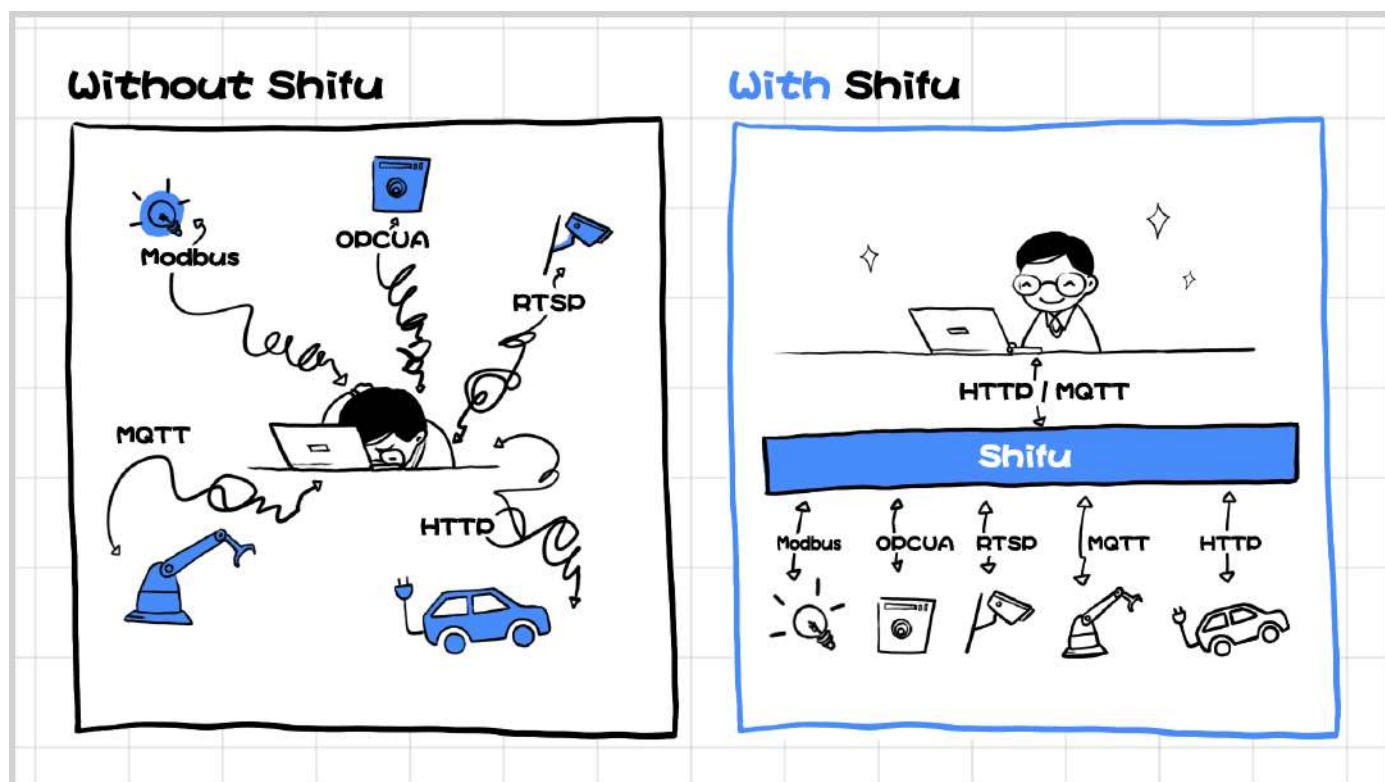
htslib是一个统一的 VCF等常见文件格式，用于高通量测序数据，是samtools和bcftools使用的核心库。C 库的实现，用于访问SAM、CRAM 和

## 12、[bcftools](#)

bcftools是一组用于变异获取和操作VCF和BCF的实用工具集合，参数众多，功能也十分强大。

## 13、[shifu](#)

一款云原生物联网开发框架。这是一个生产级别的物联网平台，它可以将物联网(IoT)设备，封装成 K8s 的最小的可部署的计算单元(pod)，直接将设备的能力和通过 API 开放出来，让物联网应用的开发变得更加简单。



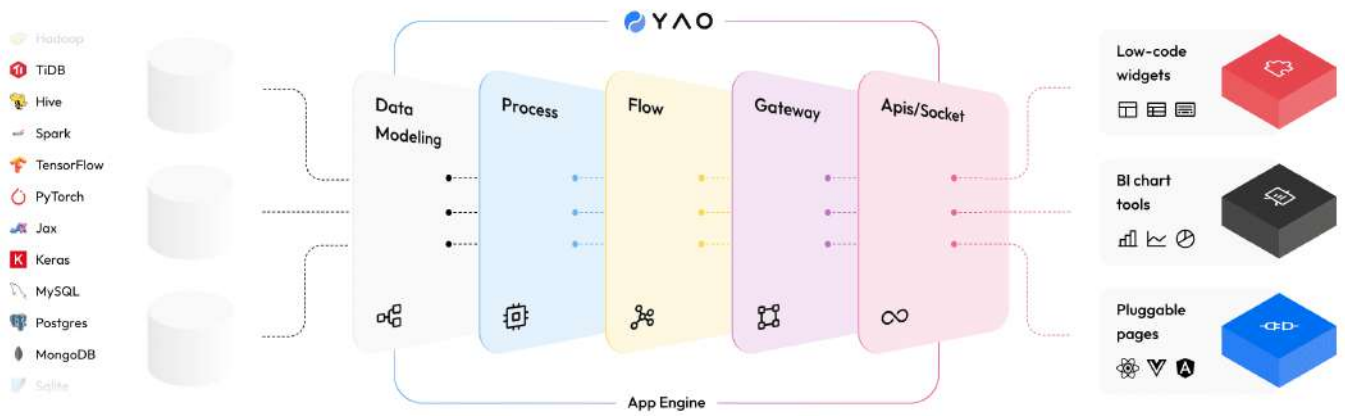
## 14、[underscore](#)

强大的 JavaScript 函数库。该库提供了 100 多个实用的函数，包括常用的 map、filter、reduce、invoke 以及更专业的辅助函数，比如函数绑定、JavaScript 模板功能、创建快速索引等，让我们可以更加方便地在 JavaScript 中实现函数式编程。

## 15、[yao](#)

一个基于Go的高性能应用程序引擎，可在几分钟内创建 Web 服务和应用程序。适用于人工智能、物联网、工业互联网、联网车辆、DevOps、能源、金融和许多其他用例。





## 16、[imglib](#)

一个轻量级的 JAVA 图像处理库，简化对图像的常见处理。

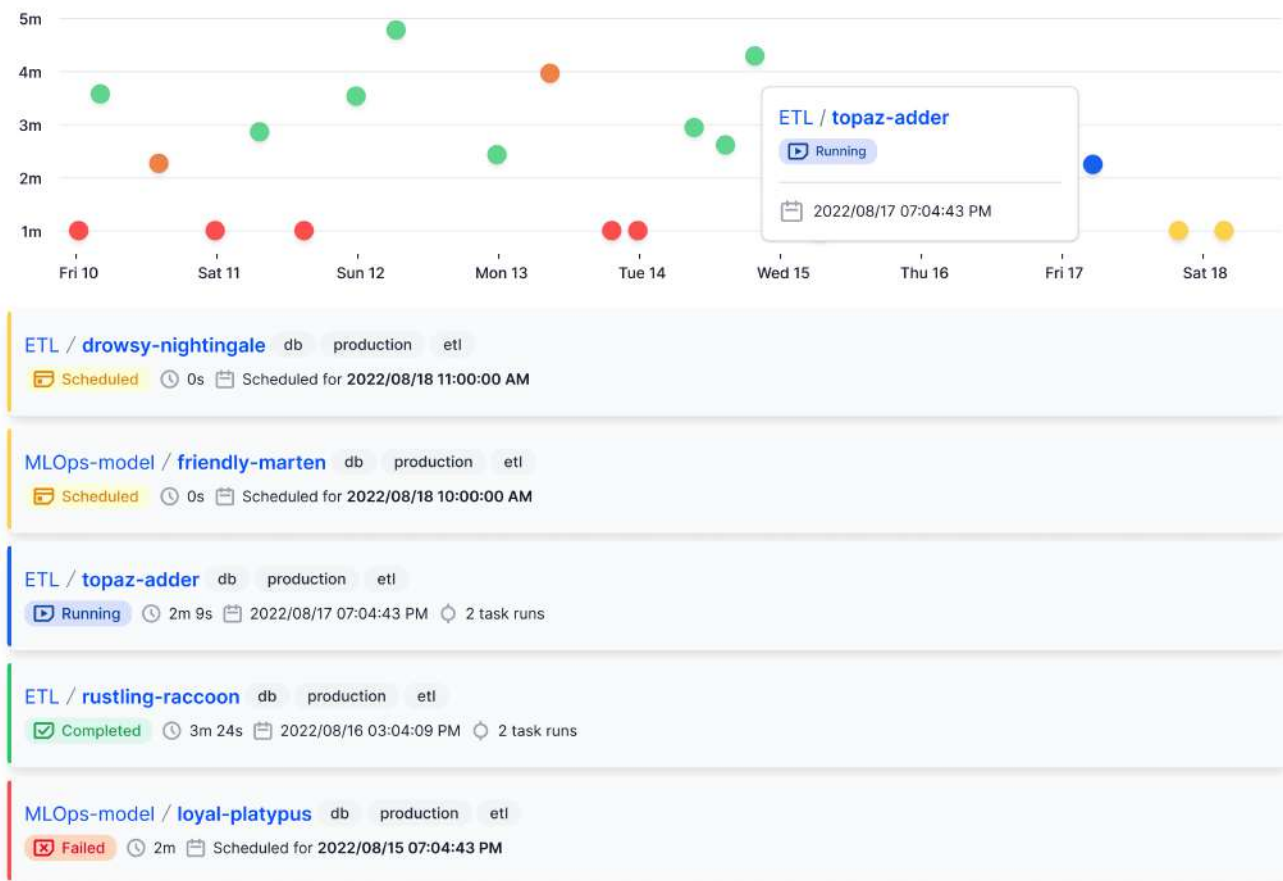
## 17、[harbor](#)

Harbor是一个用于存储和分发Docker镜像的企业级Registry服务器，可以用来构建企业内部的Docker镜像仓库。

## 18、[prefect](#)

Python 的数据流编排平台。如果将获取、清洗、处理数据的程序当作一个个分散的任务，该项目可以将这些任务整合到 workflows 中，实现在一个 Web 平台部署、安排和监控它们的执行。

### Flow Runs



## 19、[FastCFS](#)

FastCFS是一个通用的分布式文件系统，具有强一致性、高性能、高可用性，支持百亿级海量文件。FastCFS可以作为数据库、K8S、KVM、FTP、SMB、NFS的后端存储。

## 20、[pocketpy](#)

为嵌入游戏引擎而设计的 Python 解释器。一个 C++ 实现的轻量级的 Python 解释器，包含一个编译器和基于字节的虚拟机，以及交互式命令窗的实现。所有功能均集成在单个头文件 pocketpy.h 中，不包含外部依赖项可以很方便地嵌入应用，立刻拥有执行 Python 代码的能力。

## 21、[Catppuccin](#)

一套配色主题，特点是色彩柔和缤纷，不太亮不太暗，对于各种主要的应用软件都进行了适配。

## 22、[camunda-bpm-platform](#)

Camunda Platform 7 是一个灵活的工作流和过程自动化框架。它的核心是在 Java 虚拟机中运行的原生 BPMN 2.0 流程引擎。它可以嵌入到任何 Java 应用程序和任何运行时容器中。

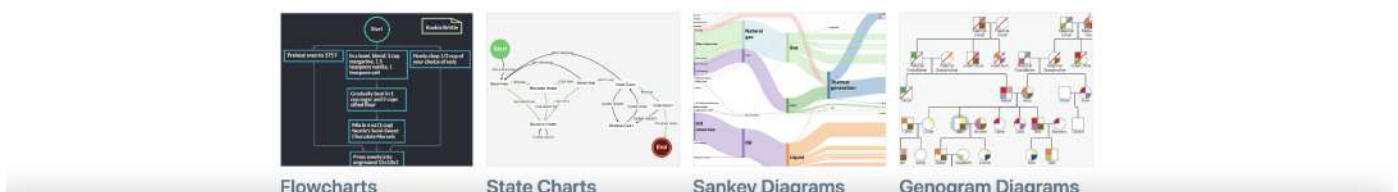
## 23、[GoJS](#)

用于交互式流程图、组织结构图、设计工具、规划工具、可视化语言的 JavaScript 图表库。



January 2023: GoJS 2.3 Released – Featuring a new SVG rendering context. See the [changelog](#) for more details.

### Build Diagrams for the Web in JavaScript and TypeScript



## 24、[docx](#)

借助漂亮的声明式 API，使用 JS/TS 轻松生成 .docx 文件。适用于 Node 和浏览器。

## 25、[scikit-opt](#)

使用 Python 封装了数学建模的常用算法，包括了遗传算法、粒子群算法、模拟退火、蚁群算法、免疫优化算法、人工鱼群算法。

文档

安装

特性

特性1: UDF (用户自定义算子)

特性2: 断点继续运行

特性3: 4种加速方法

特性4: GPU 加速

快速开始

1. 差分进化算法

2. 遗传算法

3. 粒子群算法

4. 模拟退火算法

5. 蚁群算法

6. 免疫优化算法

7. 人工鱼群算法

使用本项目的论文

参数说明

更多遗传算法

更多粒子群算法

更多模拟退火算法

遗传算法做曲线拟合

提升速度

# scikit-opt

pyPI v0.6.6 release v0.6.5 build passing codecov 92% downloads 1.16k Stars 3.9k Forks 856 chat on github

一个封装了7种启发式算法的 Python 代码库  
(差分进化算法、遗传算法、粒子群算法、模拟退火算法、蚁群算法、鱼群算法、免疫优化算法)

## 安装

```
pip install scikit-opt
```

或者直接把源代码中的 `sko` 文件夹下载下来放本地也调用可以

## 特性

### 特性1: UDF (用户自定义算子)

举例来说, 你想出一种新的“选择算子”, 如下 -> Demo code: [examples/demo\\_ga\\_udf.py#s1](#)

```
# step1: define your own operator:
def selection_tournament(algorithm, tourn_size):
    FitV = algorithm.FitV
    sel_index = []
    for i in range(algorithm.size_pop):
        aspirants_index = np.random.choice(range(algorithm.size_pop), size=tourn_size)
        sel_index.append(max(aspirants_index, key=lambda i: FitV[i]))
    algorithm.Chrom = algorithm.Chrom[sel_index, :] # next generation
    return algorithm.Chrom
```

## 26、[ChatGLM-6B](#)

ChatGLM-6B 是一个开源的、支持中英双语的对话语言模型，基于 [General Language Model \(GLM\)](#) 架构，具有 62 亿参数。结合模型量化技术，用户可以在消费级的显卡上进行本地部署（INT4 量化级别下最低只需 6GB 显存）。ChatGLM-6B 使用了和 ChatGPT 相似的技术，针对中文问答和对话进行了优化。

## 27、[skopeo](#)

能够管理远程仓库的容器镜像的工具。它能够查看远程仓库的容器镜像信息，以及执行复制、同步、删除等操作，支持 docker.io、quay.io、私有仓库等。

## 28、[secretflow](#)

蚂蚁开源的隐私计算框架。隐私计算即通过技术的手段实现数据在参与方可用不可见，让数据在安全和不泄露隐私的情况下流通、开放。该项目采用 Python 语言编写，支持包括 MPC、FL、TEE、HE、DP 在内的多种主流隐私计算技术。



## 29、[qinglong](#)

支持多种脚本语言的定时任务管理平台。这是一款定时执行脚本的平台，提供了在线管理脚本、环境变量、查看日志、秒级定时任务等功能，支持 Python3、JavaScript、shell 等脚本语言。

## 30、[llama.cpp](#)

在笔记本上运行 LLaMA 大模型。该项目实现了在 CPU 上流畅运行 LLaMA 模型，支持 macOS、Linux、Windows 操作系统。

## 31、[Chinese-LLaMA-Alpaca](#)

中文LLaMA&Alpaca大语言模型+本地部署 (Chinese LLaMA & Alpaca LLMs)

```
llama.cpp — main • serve.sh — 90x27

,
sampling: temp = 0.200000, top_k = 40, top_p = 0.950000, repeat_last_n = 64, repeat_penalt
y = 1.100000
generate: n_ctx = 2048, n_batch = 8, n_predict = 256, n_keep = 0

== Running in interactive mode. ==
- Press Ctrl+C to interject at any time.
- Press Return to return control to LLaMa.
- If you want to submit another line, end your input in '\'.

Below is an instruction that describes a task. Write a response that appropriately comple
tes the request.
> 你现在是一位导游，请介绍一下中国的首都吧
北京是中国的首都，位于中国北部，是一个历史悠久、文化底蕴深厚的城市。
> 再详细一点
北京作为中国历史上最古老的首都之一，有着丰富的历史和文化遗产，如紫禁城、故宫、天坛等知名景
点。此外，北京也是一个充满活力和创新的城市，拥有许多现代化的建筑和文化活动，如王府井大街、
798艺术区等。
> 还有别的名胜古迹吗？
当然有！北京还有很多名胜古迹，如颐和园、明月宫、北海公园等等，这些景点都具有独特的历史背景
和文化底蕴，是值得一游的景点。
> 推荐一些当地小吃吧
北京的当地小吃有烤鸭、豆汁、炸酱面、糖果等等，都是非常地道的美食，可以品尝到当地人的生活气
息。
> 
```

## 32、[FastChat](#)

一个用于训练、服务和评估基于大型语言模型的聊天机器人的开放平台。

# 五、工具软件

## 1、[zlib-searcher](#) (开源)

一个 Z-Library 的搜索器，可以自行搭建服务，就可以在上面检索和下载书籍了，示例参考 [Demo1](#)、[Demo2](#)。

需要有 Z-Library 书库的索引文件才可实现检索：<https://github.com/zlib-searcher/index>

需要配置IPFS网关，才可实现书籍资源下载，这个IPFS网关可以自部署，也可以用别人搭建好的。

```
https://cloudflare-ipfs.com
https://dweb.link
https://ipfs.io
https://dw.oho.im
```

🔍 西游记  @ 作者  出版社

📄 扩展名  🗣️ 语言  # ISBN

🔍 复杂搜索

书名	作者	出版社	扩展名	文件大小	语言	年份	页数
西游记	(明) 吴承恩		mobi	2.13 MB	Other	1998	

西游记	
ID: 16532211	IPFS CID: bafykbzaceb5dit2kmnkopxgzvcstuu4dwqw2zy5ewijmr7xzmsac5o4rgdou
作者: (明) 吴承恩	出版社: 未知
语言: Other	年份: 1998
	扩展名: mobi
	文件大小: 2.13 MB
	页数: 未知
	ISBN: 未知

<https://cloudflare-ipfs.com>
<https://dweb.link>
<https://ipfs.io>
<https://dw.oho.im>

西游记	吴承恩 & chenjin5.com	cj5	mobi	2.06 MB	Other	2004
西游记	吴承恩	银河系出版社_02	mobi	2.13 MB	Other	2015
西游记	吴承恩	epub掌上书苑	epub	1.2 MB	Other	2010
西游记	吴承恩 & chenjin5.com	chenjin5.com 沉金书屋海量...	mobi	2.04 MB	Chinese	2011
西游记	吴承恩 & chenjin5.com	chenjin5.com 海量电子书免...	mobi	2.05 MB	Chinese	2021

## 2、github-readme-stats (开源)

为Github自述文件动态生成的统计信息。

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

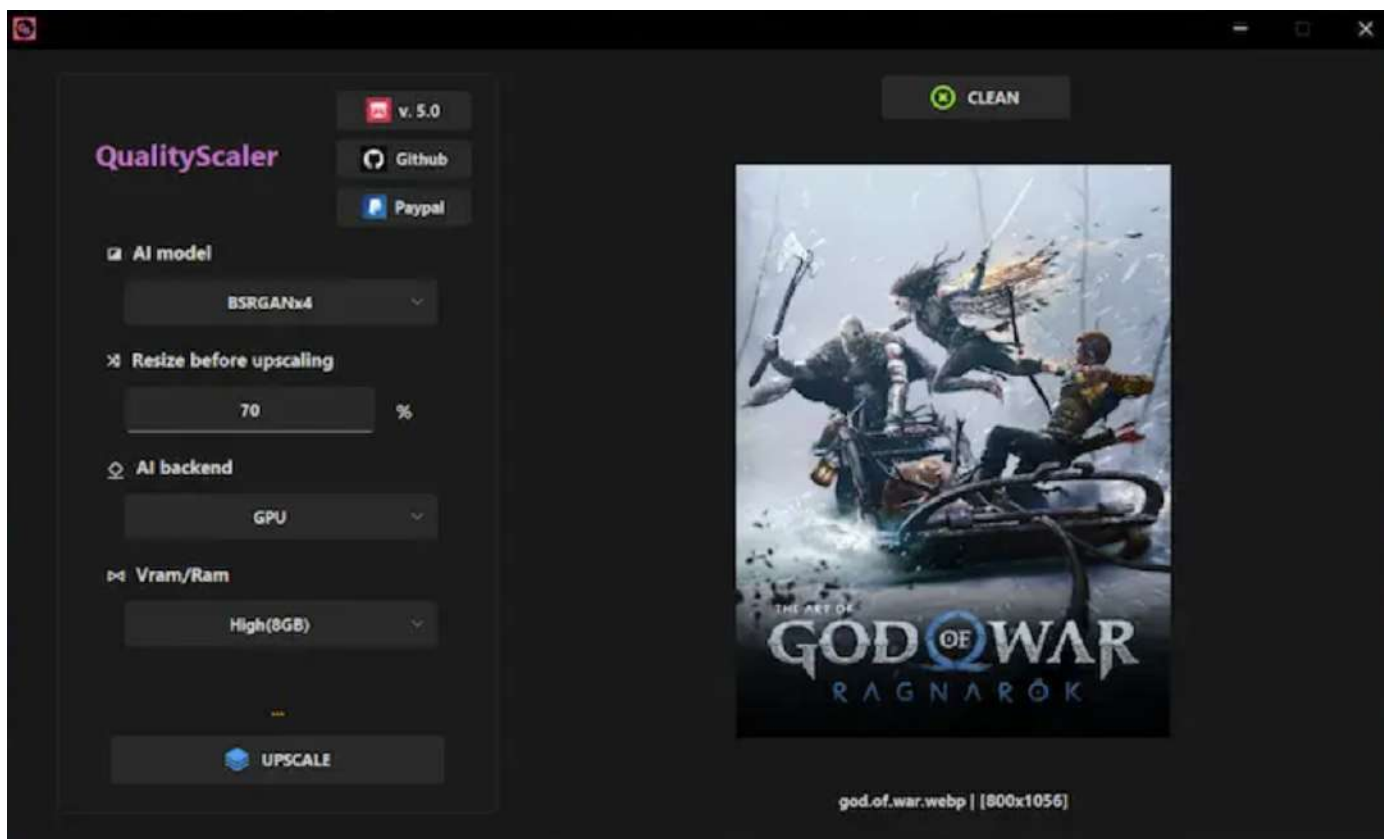
**Anurag Hazra's GitHub Stats**

- ☆ Total Stars: 2.2k
- 🕒 Total Commits: 1k
- 🔗 Total PRs: 202
- 🔔 Total Issues: 95
- 📁 Contributed to: 63

**A++**

### 3、[QualityScaler](#) (开源)

一个 Windows 桌面软件，可以提升图片和视频的分辨率，让它们变得更清晰。



### 4、[grpcurl](#) (开源)

类似 cURL 但用于 gRPC 的工具。一款实现与 gRPC 服务器交互的命令行工具，可以轻松请求 gRPC 服务，就像 gRPC 版的 cURL 一样好用。

```
# 安装
brew install grpcurl
# 使用
grpcurl grpc.server.com:443 my.custom.server.Service/Method
```

### 5、[HummerRisk](#) (开源)

云原生安全检测平台。该项目用非侵入的方式，解决云原生环境的安全和治理问题。支持主流公/私有云资源的安全检测、漏洞扫描、一键获取报告等功能。

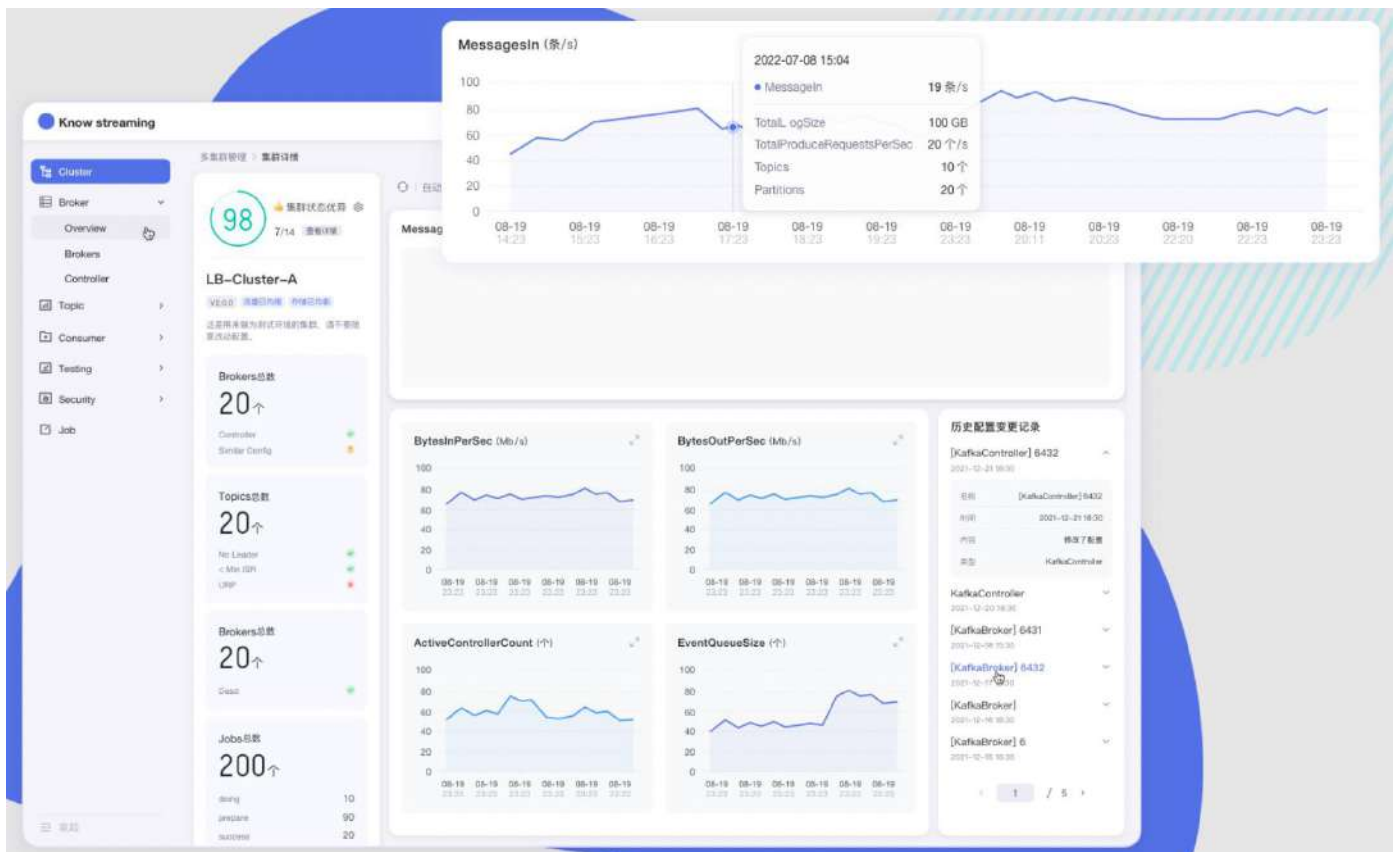


## 6、bandit (开源)

查找 Python 代码中常见安全问题的工具。该项目是 PyCQA 出品的 Python 代码检测工具。

## 7、KnowStreaming (开源)

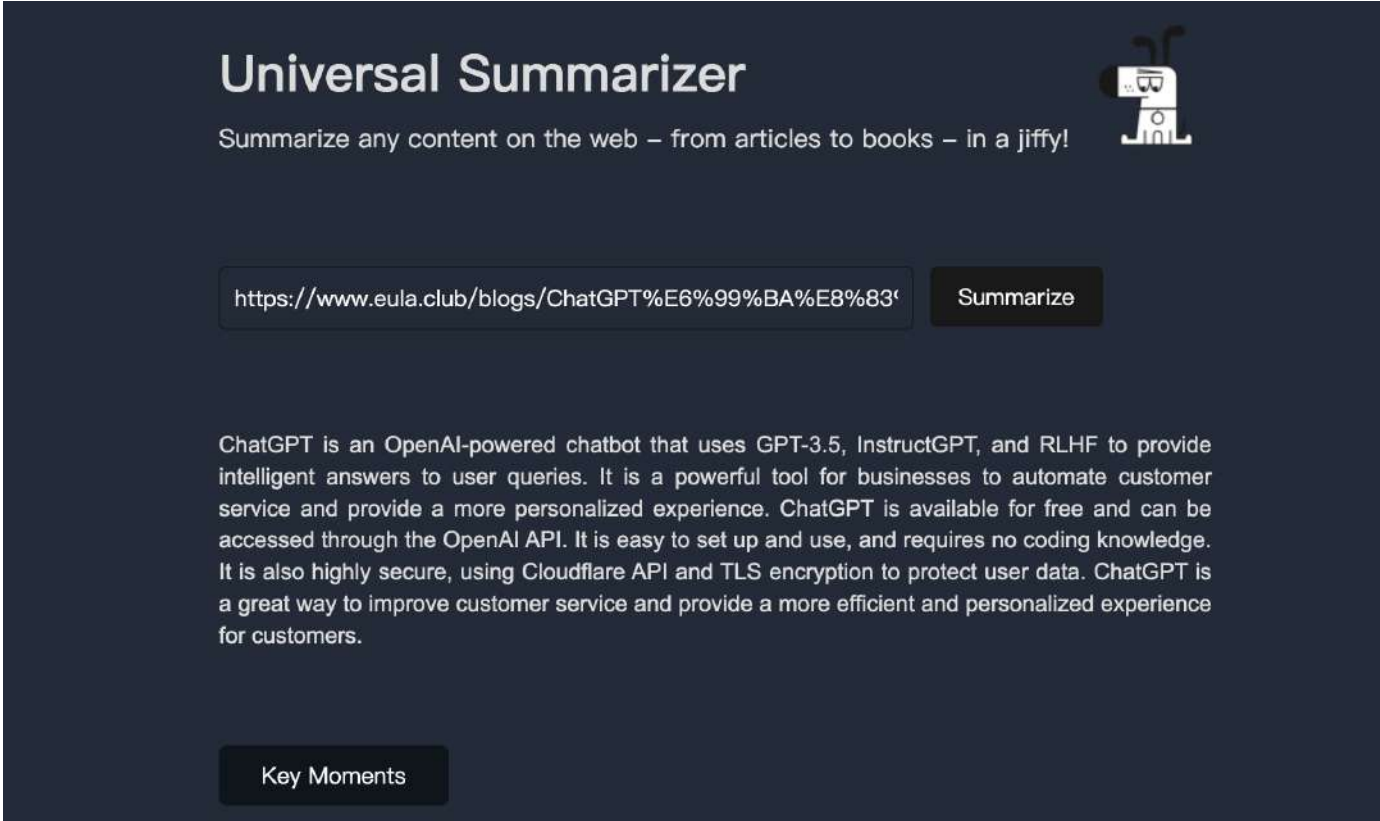
Know Streaming是一套云原生的Kafka管控平台，脱胎于众多互联网内部多年的Kafka运营实践经验，专注于Kafka运维管控、监控告警、资源治理、多活容灾等核心场景。在用户体验、监控、运维管控上进行了平台化、可视化、智能化的建设，提供一系列特色的功能，极大地方便了用户和运维人员的日常使用。





## 8、[文章总结器](#) (免费)

搜索引擎 Kagi 推出的一个工具，输入网址，就能快速总结出该网址文章的大意，支持中文网址。



The screenshot shows the 'Universal Summarizer' interface. At the top, the title 'Universal Summarizer' is displayed in white on a dark blue background. Below the title is the tagline 'Summarize any content on the web – from articles to books – in a jiffy!' and a small robot icon. A text input field contains the URL 'https://www.eula.club/blogs/ChatGPT%E6%99%BA%E8%83', and a 'Summarize' button is positioned to its right. Below the input field, a paragraph of text describes ChatGPT as an OpenAI-powered chatbot. At the bottom of the interface, there is a 'Key Moments' button.

# Universal Summarizer

Summarize any content on the web – from articles to books – in a jiffy!

<https://www.eula.club/blogs/ChatGPT%E6%99%BA%E8%83> Summarize

ChatGPT is an OpenAI-powered chatbot that uses GPT-3.5, InstructGPT, and RLHF to provide intelligent answers to user queries. It is a powerful tool for businesses to automate customer service and provide a more personalized experience. ChatGPT is available for free and can be accessed through the OpenAI API. It is easy to set up and use, and requires no coding knowledge. It is also highly secure, using Cloudflare API and TLS encryption to protect user data. ChatGPT is a great way to improve customer service and provide a more efficient and personalized experience for customers.

Key Moments

## 9、[帮小忙](#) (免费)

腾讯推出的网页小工具的集合网站。



## 10、[TrollStore](#) (开源)

TrollStore 是一个永久签名的非越狱 App。通过它，我们可以安装任意未签名的 IPA 文件。

它基于一个存在于 iOS 14.0 -- iOS 15.4.1 当中的 AMFI/CoreTrust 漏洞：iOS 不验证用于签署二进制文件的根证书是否合法。因此，如果你的IOS设备系统版本在这个范围内，则可以安装 TrollStore，而不论是否越狱。

版本/设备	arm64 (A8 - A11)	arm64e (A12 - A15, M1)
13.7及以下	不支持 (CT Bug 仅在 14.0 中引入)	不支持 (CT Bug 仅在 14.0 中引入)
14.0 - 14.8.1	<a href="#">checkra1n + TrollHelper</a>	<a href="#">TrollHelperOTA (arm64e)</a>
15.0 - 15.4.1	<a href="#">TrollHelperOTA (iOS 15+)</a>	<a href="#">TrollHelperOTA (iOS 15+)</a>
15.5 测试版 1 - 4	<a href="#">TrollHelperOTA (iOS 15+)</a>	<a href="#">TrollHelperOTA (iOS 15+)</a>
15.5 (RC)	不支持 (已修复 CT 错误)	不支持 (已修复 CT 错误)
15.6 测试版 1 - 5	<a href="#">SSH 虚拟磁盘</a>	<a href="#">TrollHelperOTA (arm64e)</a>
15.6 (RC1/2) 及以上	不支持 (已修复 CT 错误)	不支持 (已修复 CT 错误)

## 11、[ambie](#) (开源)

Windows 上的白噪声应用。一款播放白噪声和自然声音的应用，比如下雨、海滩等声音，支持混合、在线下载声音和专注功能。工作时使用可以帮助你集中注意力，还能在放松时使用有助于睡眠。



## 12、[NightVision](#) (开源)

夜视仪iOS应用，使用 iPhone 的激光镜头，在完全无光环境下扫描和检测距离。

## 13、[sigma-file-manager](#) (开源)

一款先进的文件管理器。这是一款免费的文件管理器，由开源社区维护。支持智能搜索、自定义主页、文件共享、文件下载、智能拖放、文件保护等功能，适用于 Windows 和 Linux。



#### 14、[gsudo](#) (开源)

适用于 Windows 的 sudo 命令行工具。它是 Windows 上的 sudo，允许用户以最高权限运行命令，拥有与 Unix/Linux sudo 类似的使用体验，支持 CMD、PowerShell、git-bash 等。

#### 15、[sqlc](#) (开源)

将 SQL 转成类型安全的 Go 代码的工具。它可以将输入的 SQL 语句，自动转化成类型安全、可读的操作数据库的 Go 代码，支持 MySQL、PostgreSQL 和 SQLite 数据库。

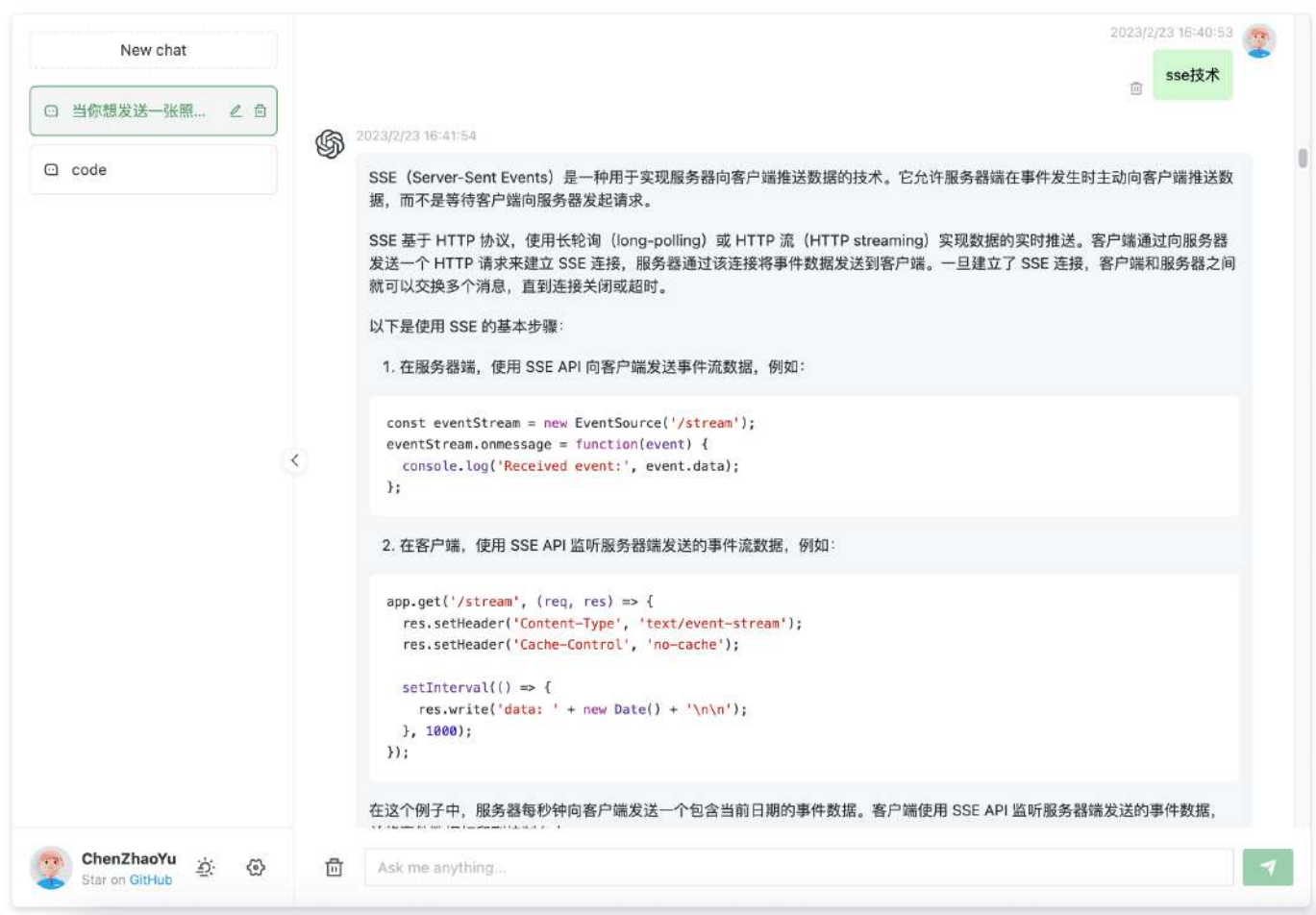
```
INPUT query.sql sqlc.json sqlc-version.txt OUTPUT db/db.go db/models.go db/query.sql.go
1 -- Example queries for sqlc
2 CREATE TABLE authors (
3 id BIGSERIAL PRIMARY KEY,
4 name text NOT NULL,
5 bio text
6 );
7
8 -- name: GetAuthor :one
9 SELECT * FROM authors
10 WHERE id = $1 LIMIT 1;
11
12 -- name: ListAuthors :many
13 SELECT * FROM authors
14 ORDER BY name;
15
16 -- name: CreateAuthor :one
17 INSERT INTO authors (
18 name, bio
19 ) VALUES (
20 $1, $2
21 )
22 RETURNING *;
23
24 -- name: DeleteAuthor :exec
25 DELETE FROM authors
26 WHERE id = $1;
1 // Code generated by sqlc. DO NOT EDIT.
2 // source: query.sql
3
4 package db
5
6 import (
7 "context"
8 "database/sql"
9 )
10
11 const createAuthor = `-- name: CreateAuthor :one
12 INSERT INTO authors (
13 name, bio
14 ) VALUES (
15 $1, $2
16 )
17 RETURNING id, name, bio
18 `
19
20 type CreateAuthorParams struct {
21 Name string
22 Bio sql.NullString
23 }
24
25 func (q *Queries) CreateAuthor(ctx context.Context, arg CreateAuthorParams) (Author, error) {
26 row := q.db.QueryRowContext(ctx, createAuthor, arg.Name, arg.Bio)
27 var i Author
28 err := row.Scan(&i.ID, &i.Name, &i.Bio)
29 return i, err
30 }
31
```

#### 16、[ghz](#) (开源)

一款用 Go 开发的专门用来压测 gRPC 服务的命令行工具，它使用简单、高效、支持自定义参数。

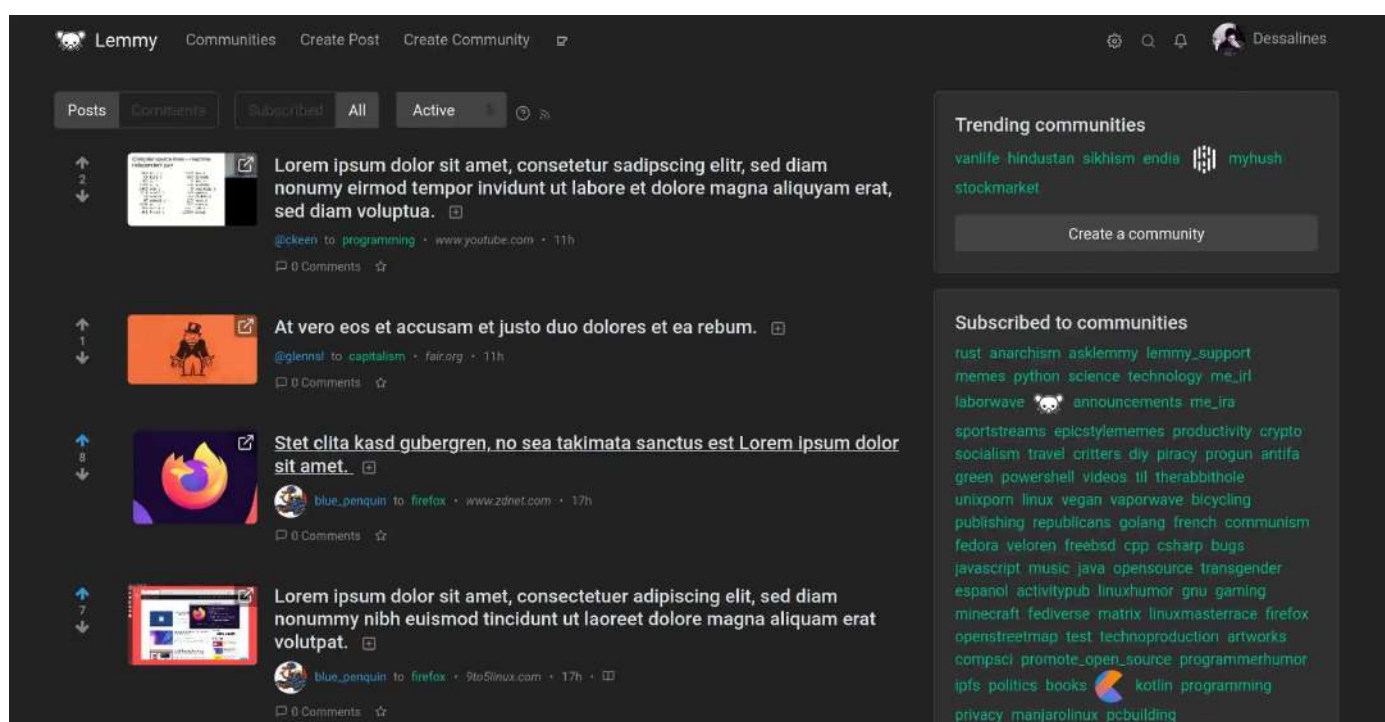
#### 17、[chatgpt-web](#) (开源)

一款可自定义 API 的 ChatGPT 演示网页。基于 Express 和 Vue3 构建的 GPT-3 模型演示网页，支持接入 GPT-3 API 或网页 ChatGPT。



## 18、lemmy (开源)

Rust 写的链接聚合论坛。该项目基于 Rust 的 Web 框架 Actix 和 Diesel ORM 库构建，它是一个类似 Hacker News 的网站，用户可以在上面订阅感兴趣的话题、发布链接、讨论和投票。



## 19、[onefetch](#) (开源)

查看 Git 仓库信息的命令行工具。一款由 Rust 编写的命令行查看 Git 信息的工具，它可以直接在终端中展示本地 Git 仓库的详细信息，比如开源协议、提交次数、代码统计等信息。



```
o2sh ~ git version 2.37.1
-----
Project: electron (345 branches, 2037 tags)
Description: Build cross platform desktop apps
             with JavaScript, HTML, and CSS
HEAD: 75d2caf451 (main, origin/main)
Pending: 1+- 1+
Version: v23.0.0-nightly.20221111
Created: 9 years ago
Languages: ██████████
            ● C++ (49.6 %) ● TypeScript (35.0 %)
            ● JavaScript (7.3 %) ● HTML (5.3 %)
            ● Python (2.6 %) ● CSS (0.1 %)
            ● Other (0.1 %)
Dependencies: 71 (Npm)
Authors: 33% Cheng Zhao 8454
         12% Kevin Sawicki 2998
         4% Shelley Vohr 907
Last change: a day ago
Contributors: 1431
Repo: https://github.com/electron/electron
Commits: 25378
Lines of code: 115369
Size: 18.08 MiB (2328 files)
License: MIT

o2sh ~/dev/electron █
```

## 20、[cog](#) (开源)

将机器学习模型打包到容器的工具。可通过配置将机器学习模型所需的环境和依赖，自动打包到容器里方便部署，让你不再为编写 Docker 文件和 CUDA 而痛苦，还能自动启动 HTTP 接口服务方便调用。

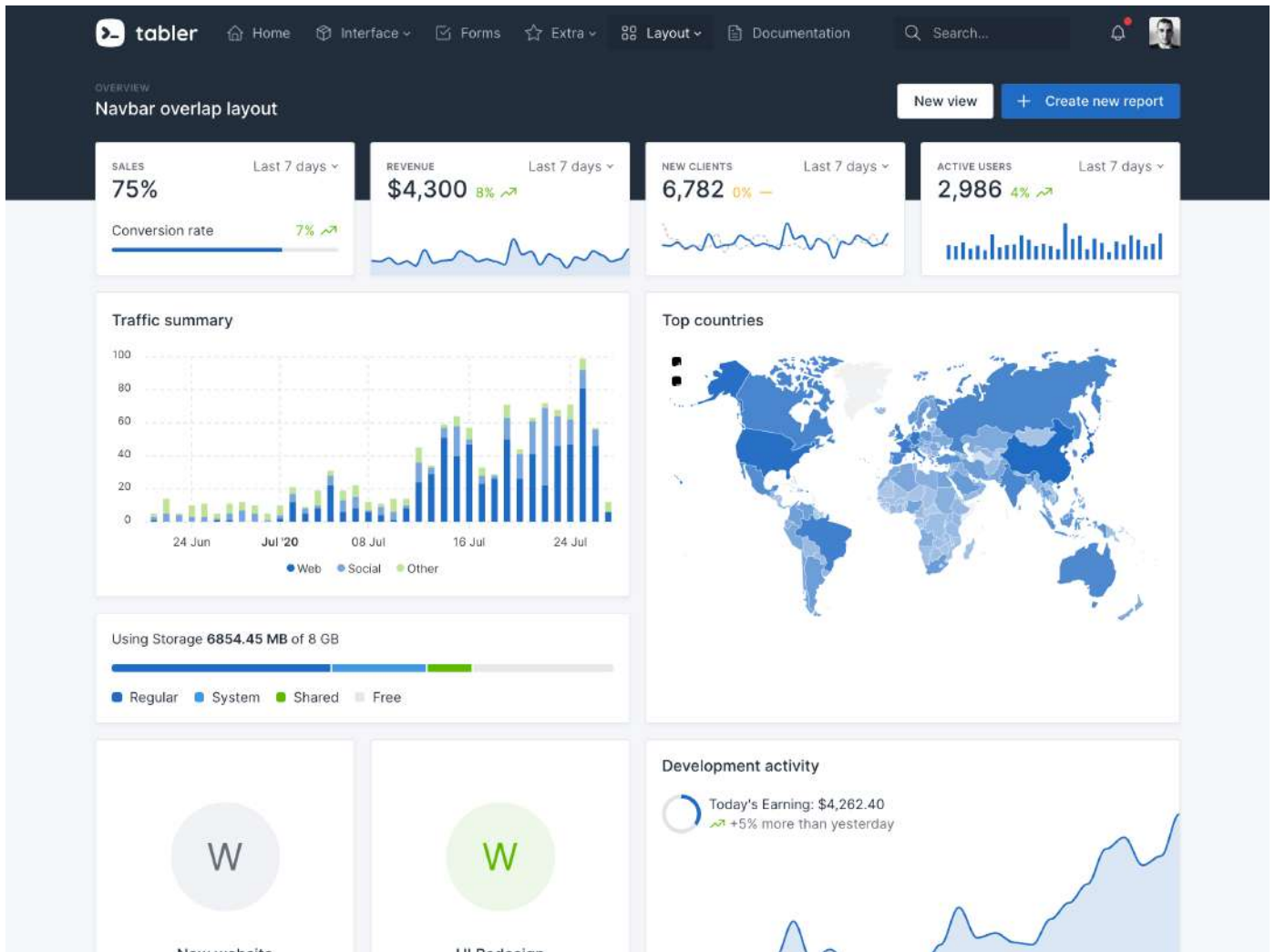
## 21、[幕境](#) (开源)

英语学习软件，用户上传原版视频和字幕，软件会自动生成词库。以后播放视频时，遇到词库里面的单词，该词就会以弹幕形式自动展示释义。



## 22、[Tablet](#) (开源)

一个基于 Bootstrap 的网页 UI 框架，专用于开发管理后台。



## 23、[bilingual\\_book\\_maker](#) (开源)

用 AI 翻译来制作 epub 图书的工具，默认用了 ChatGPT 模型，可用 `--model gpt3` 来使用 GPT-3 模型。  
`bilingual_book_maker` 支持所有语言，可通过 `--language "Simplified Chinese"` 来设定翻译语言。如果你在  
读一本外文图书，可以用 `bilingual_book_maker` 翻译整书为你熟悉的语言。



我們居住的那個城鎮，是典型的大都會郊區的中產階級住宅區。對在那里居住期間擁有一定親熟關係的同學，他們大多都住在整潔的一戶屋里面。雖然屋子大小有些差別，但他們都有前廳、花園，和長滿樹木的花園。朋友的父親大多都是在公司上班或者從事專業職業。只有極少數家庭的母親有工作。凡是家裡都養一只狗或是一只貓。我在那時從未聽說會把住在公寓或是公寓裡的人。後來我搬到另一個附近的城鎮，幾乎也是同樣的情形。所以我進入大學，直到來到東京之前，我一直堅信，普通人都是穿著領帶上班，住著一個有院子的屋子，養著一隻狗或者一隻貓，我幾乎沒有感受到其他的生活方式。

大抵の家には二人か三人の子供がいた。それが僕の住んでいた世界における平均的な子供の数だった。少年時代から思春期にかけて持った何人かの友人の顔を思い浮かべてみても、彼らは一人の例外もなく、まるで判で押したみたいに二人兄弟か、あるいは三人兄弟の一員だった。彼らは二人兄弟でなければ、三人兄弟であり、三人兄弟でなければ二人兄弟だった。六人も七人も子供がいる家庭は稀だったが、一人しか子供がいない家庭というのはそれ以上に稀だった。

大概的家庭都有两到三个小孩。这也是我生活过的世界里小孩平均数量。当回想起我从少年时期到青春期间所交往过的几个朋友，没有一个例外，他们几乎都是两个兄弟，或者三个兄弟中的任何一位。他们不是双胞胎，就是三胞胎；不是三胞胎，就是双胞胎。有六七个孩子的家庭很罕见，只有一个孩子的家庭更少见。

でも僕には兄弟というものがただの一人もいなかった。僕は一人っ子だった。そして少年時代の僕はそのことでずっと引け目のようなものを感じていた。自分はこの世界にあってはいわば特殊な存在なのだ、他の人々が当然のこととして持っているものを、僕は持っていないのだ。

：可是我一个人没有兄弟姐妹，我是个独生子，小时候我就有一种被人冷当的感觉，觉的自己在这个世上是个特殊的存在，没有其他人拥有的东西。

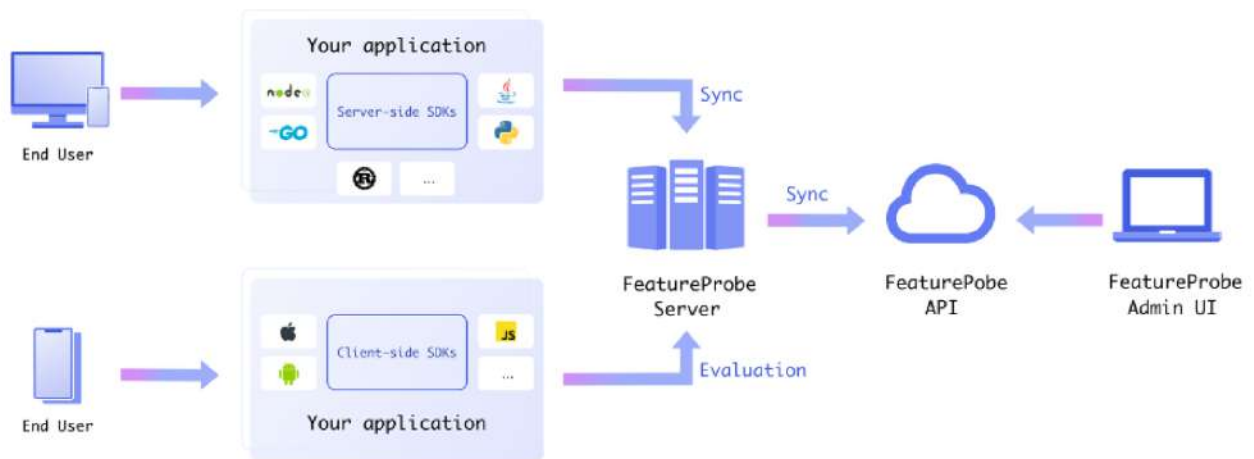
子供の頃、僕はこの「一人っ子」という言葉がいやでたまらなかった。その言葉を耳にするたびに、自分には何か欠けているのだということをあらためて思い知らされることになった。その言葉はいつも僕に向かってまっすぐに指をつきつけていた。お前は不完全なのだぞ、と。

当我还是小孩的时候，“一个人的孩子”这句话让我感到非常不舒服。每次听到这句话，我都会认真清醒地意识到自己缺少了什么。这句话总是直指我，好像在说：“你不完整哦。”

一人っ子が両親にあまやかされていて、ひ弱で、おそろしくわがままだというのは、僕が住んでいた世界においては揺るぎない定説だった。それは高い山に登れば気圧が下がるとか、雌の牛は多量の乳を出すとかいうのと同じ種類の自然の摂理とみなされていた。だから僕は誰かに兄弟の数を訊かれるのが嫌でたまらなかった。兄弟がいないと聞いただけで人々は反射的にこう思うのだ。こいつは一人っ子だから、両親にあまやかされていて、ひ弱で、お

## 24、[FeatureProbe](#) (开源)

FeatureProbe 是一个开源的“特性”管理服务，包含：灰度放量、AB 实验、实时配置变更等针对“特性粒度”的一系列管理操作。特性粒度的发布管理是实现 DevOps 的核心工具之一，通过“特性”开关，可以降低分支开发带来的合并复杂性，轻松实现主干开发以及持续交付。



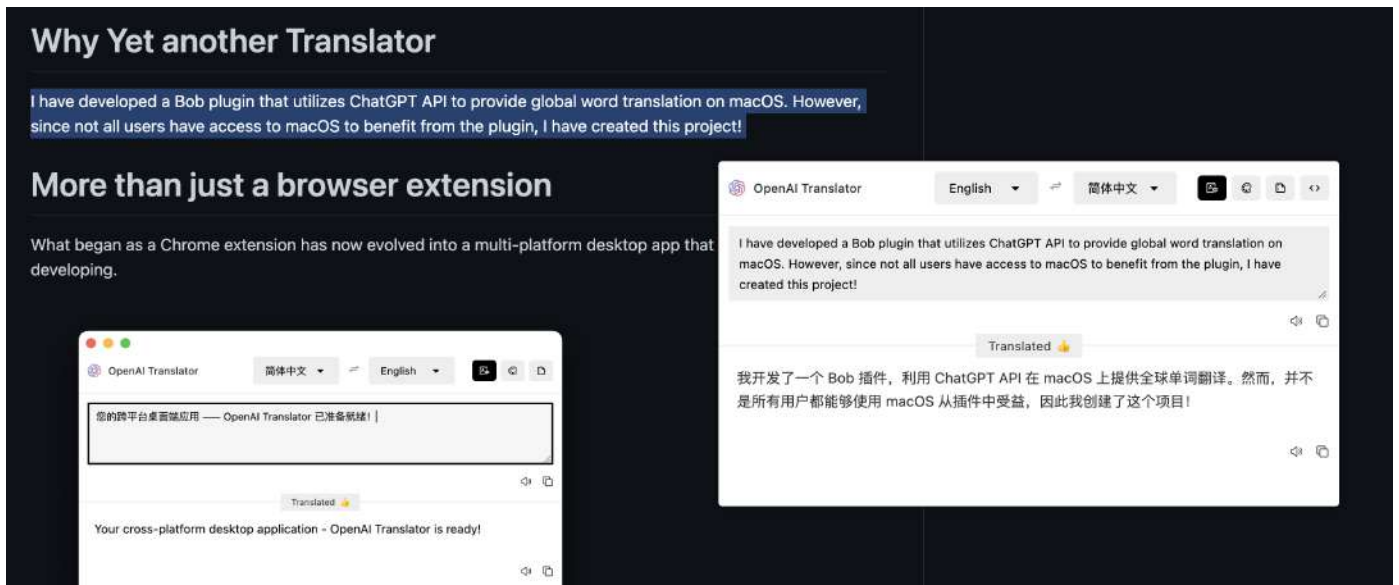
## 25、[openai-translator](#) (开源)

基于 ChatGPT API 的划词翻译浏览器插件和跨平台桌面端应用。特性：

- 支持 55 种语言的相互翻译、润色和总结功能
- 支持实时翻译、润色和总结，以最快的速度响应用户

- 支持自定义翻译文本
- 支持一键复制
- 全平台 (Windows + macOS + Linux) 支持

它的使用也非常简单，下载插件 / 应用之后，填入你的 OpenAI API Key 就能使用。



## 26、[ChatPDF](#) (免费)

基于ChatGPT的一个应用，上传PDF文档即可让它替你学习，然后你可以直接问它相关问题。ChatPDF的原理是先对上传的PDF进行分析，为文件中每个段落创建语义索引。当用户提出一个问题后，工具就会把关联语段发送给ChatGPT，然后让它结合问题进行解读。

## 27、[ChatGPT-Telegram-Workers](#) (开源)

基于JS实现，可在 Cloudflare Workers 上免 VPS 部署 ChatGPT Telegram Bot，交互按钮是中文的。

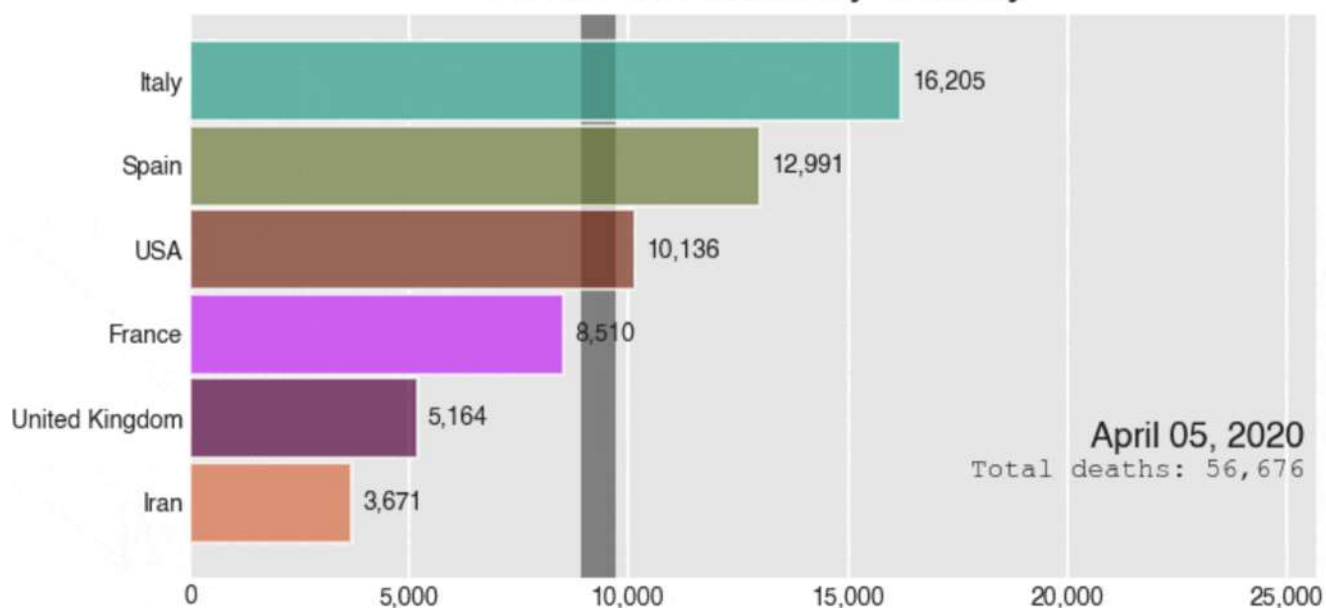
## 28、[chatgpt\\_telegram\\_bot](#) (开源)

基于Python实现，可使用 Docker 在自己的 VPS 上部署 ChatGPT Telegram Bot，交互按钮是英文的。

## 29、[bar\\_chart\\_race](#) (开源)

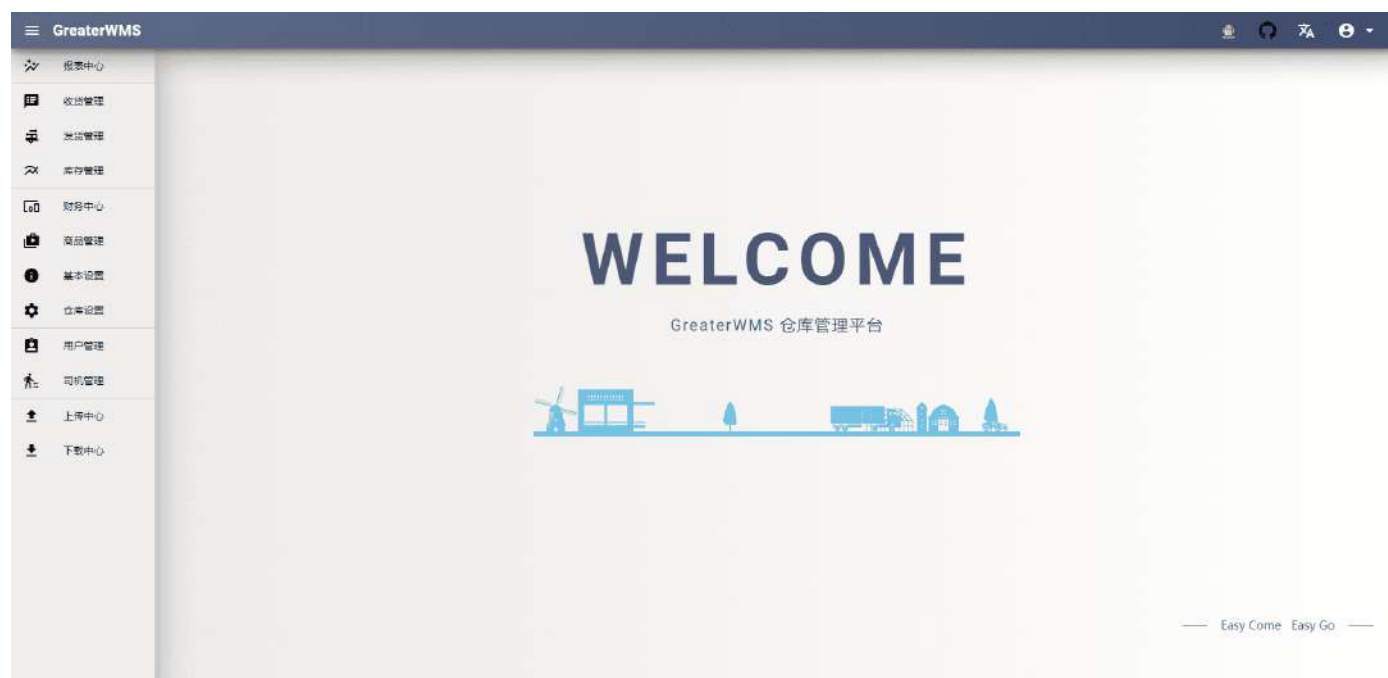
基于 Python 的动态条形图。通过该项目可以用 Python 创建条形图比赛动画，显示数据排名的动态条形图，直观地展示数据变化过程。

## COVID-19 Deaths by Country



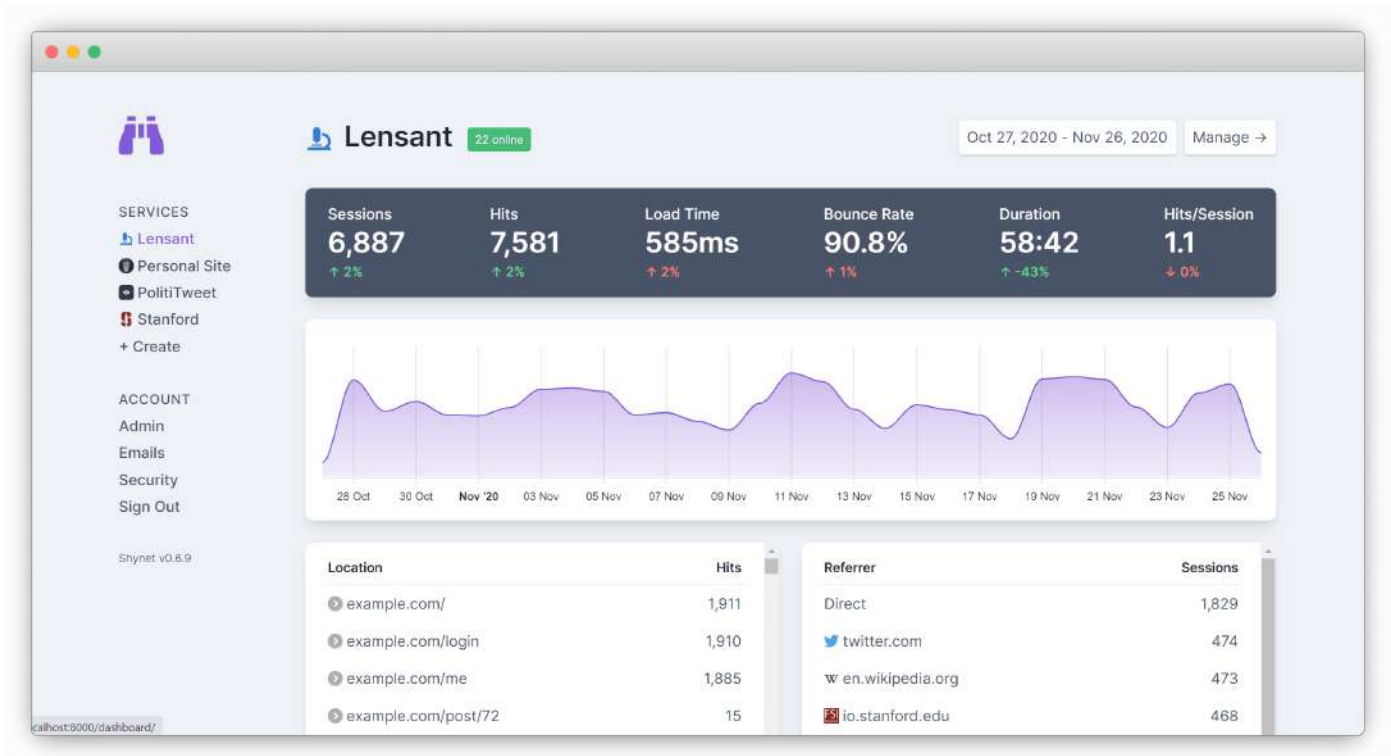
### 30、[GreaterWMS](#) (开源)

可商用的开源仓库管理系统。该项目是采用福特亚太区售后物流仓储供应链流程的仓库系统，它提供了客户管理、订单管理、库存管理、供应商管理、盘点等模块，支持手机、电脑等多种设备。



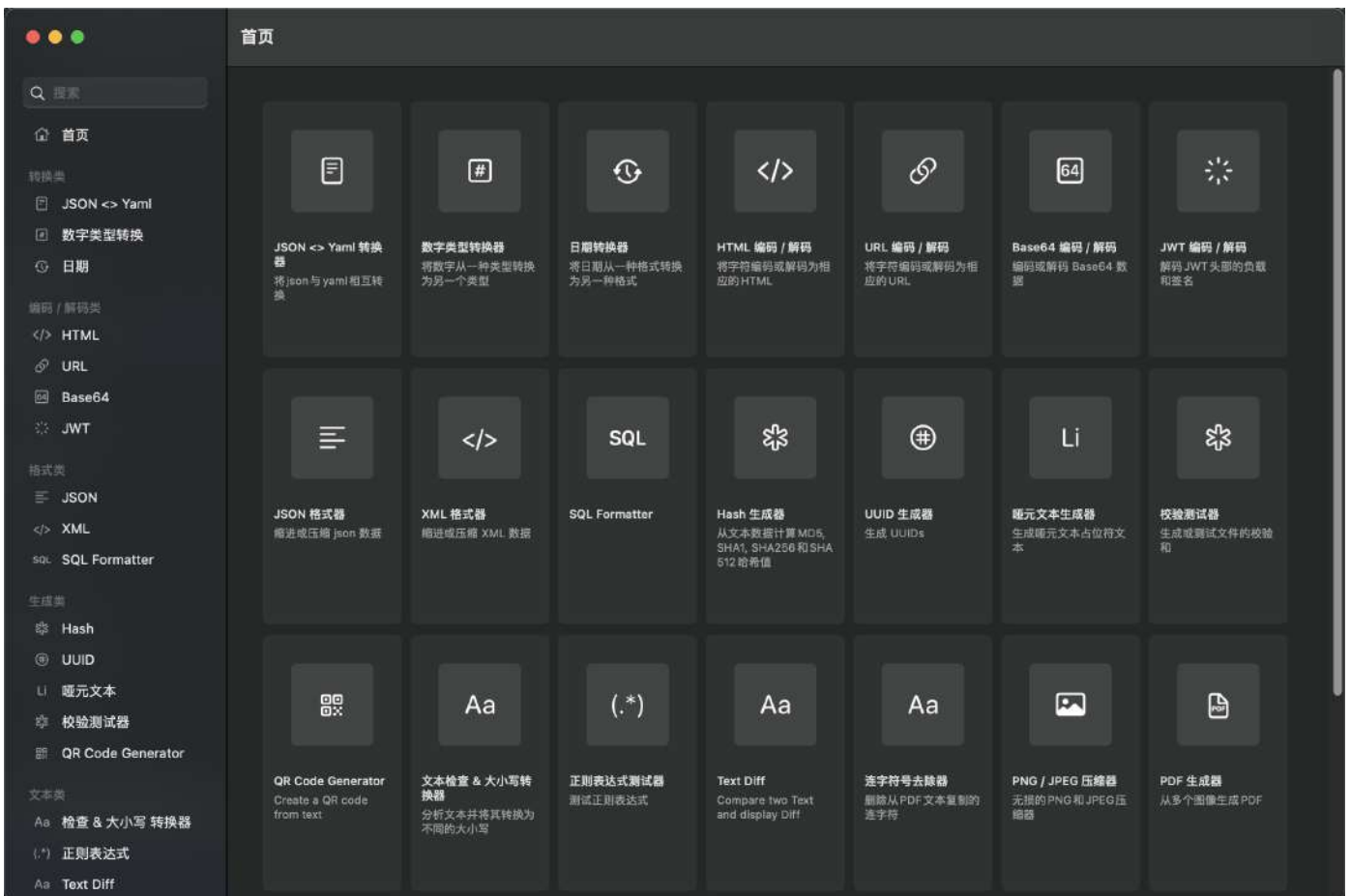
### 31、[shynet](#) (开源)

极简的网站分析平台。这是一个基于 Django 构建的网站分析平台，它很小、够用、界面友好、不追踪 cookie、支持多用户，追踪脚本不到 1KB。



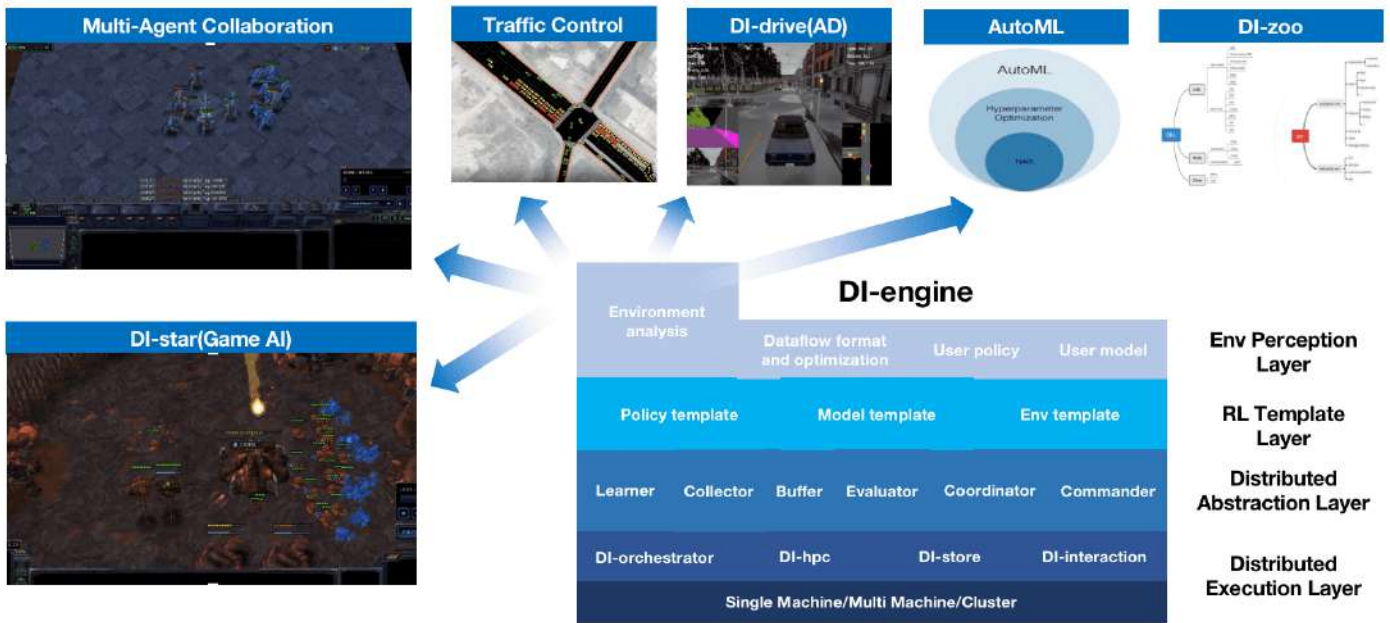
### 32、[DevToysMac](#) (开源)

该项目是 DevToys 的 macOS 版本，无需安装下载解压后即可使用。它同样实现了程序员日常开发会用到的功能，比如时间戳转化、Base64 编/解码、JSON 格式化等。



### 33、[DI-engine](#) (开源)

OpenDILab 开源的决策 AI 平台。这是一个基于 PyTorch 的通用决策智能引擎，为开发者提供了 60+ 种算法、40+ 类型环境。支持各类定制化的训练和实际决策智能应用，比如游戏 AI、自动驾驶和生物序列预测等。

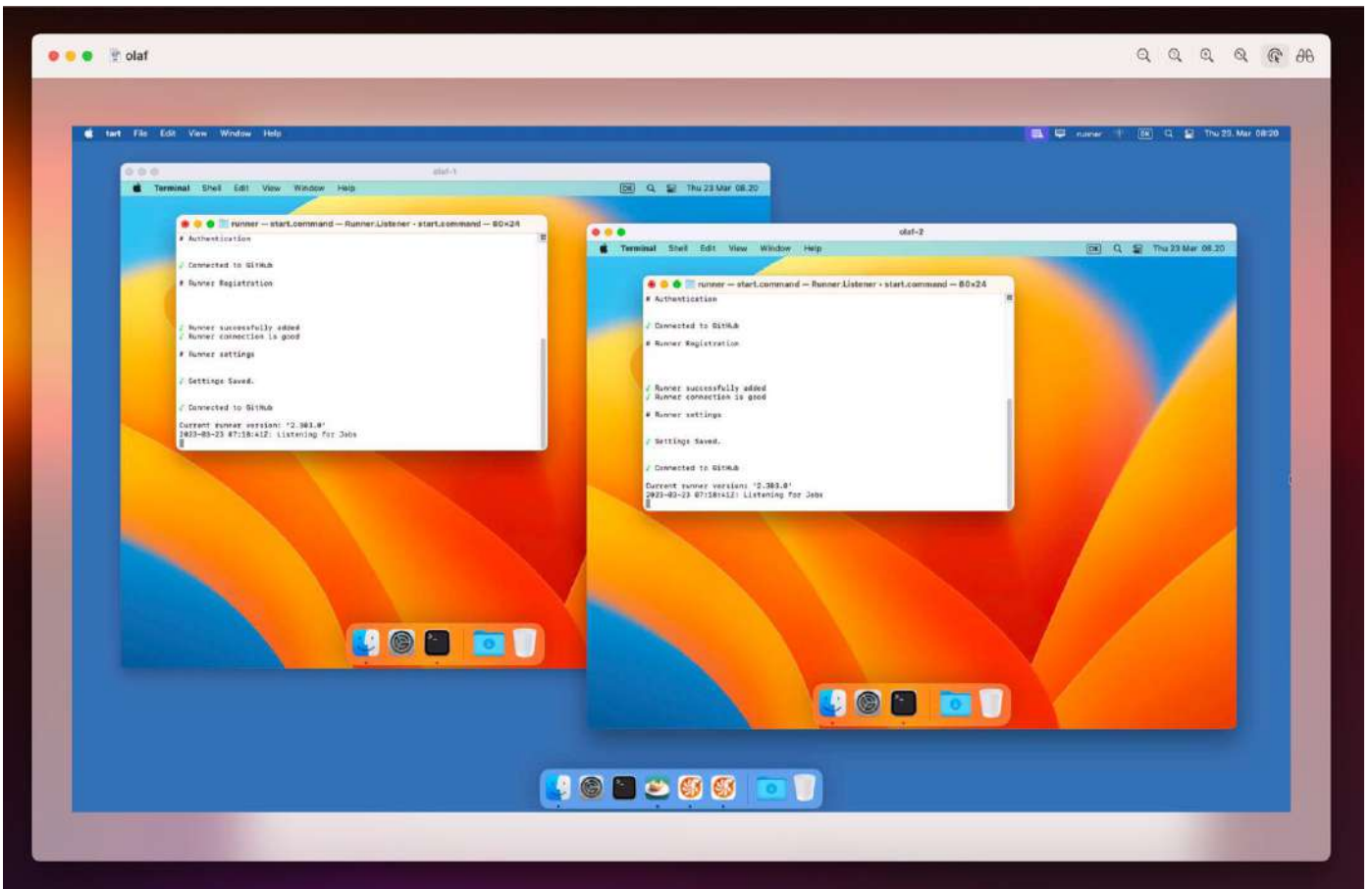


### 34、[chatgpt-retrieval-plugin](#) (开源)

ChatGPT 检索插件让您可以通过使用日常语言提问来轻松搜索和查找个人或工作文档。

### 35、[Tartelet](#) (开源)

一个在本机运行 GitHub Actions 的工具，通过虚拟机来执行任务，只支持 MacOS 系统。



### 36、[gpt4-pdf-chatbot-langchain](#) (开源)

为您的 PDF 文件创建 ChatGPT 聊天机器人，相当于ChatPDF的开源版。

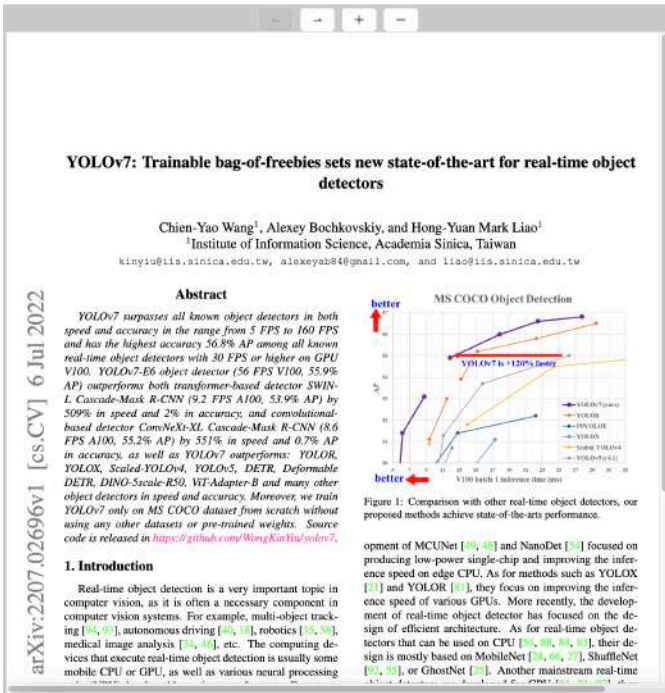
### 37、[nofwl](#) (开源)

ChatGPT 桌面应用程序（支持 Mac、Windows 和 Linux），使用 OpenAI API 密钥的方式进行访问。



### 38、[ChatPaper](#) (开源)

全流程加速科研，利用chatgpt进行论文总结+润色+审稿+审稿回复。[官方体验地址](#)



Basic Information:

- Title: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors (YOLOv7: 一种可训练的免费袋包装技术为实时目标检测设定了新的最先进技术)
- Authors: Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao
- Affiliation: Institute of Information Science, Academia Sinica, Taiwan (中央研究院資訊科學研究所)
- Keywords: YOLOv7, object detection, real-time detection, deep learning, convolutional neural network (CNN)
- Uri: <https://arxiv.org/abs/2106.09747>, Github: <https://github.com/WongKinYiu/yolov7>.

Summary:

- (1): 本篇文章介绍了实时目标检测在计算机视觉领域的键应用价值;
- (2): 本文提出的YOLOv7算法使用优化模块和方法进行训练优化以提高目标检测的精度,并解决了目标检测方法中的两个新问题;
- (3): 本文中的YOLOv7算法使用了优化的卷积, re-parameterized模块, 可靠的approximate non-maximum suppression方法, Dynamic Layer label Smoothing以及可训练的freebies方法等, 融合了group convolutions, cascade结构, 道路轨迹以及深度监督等技术;
- (4): YOLOv7-E6目标检测器在V100上达到了56FPS和55.9% AP的性能, 与SWIN-L Cascade-Mask R-CNN相比, 在速度和精度方面分别提高了509%和2%, 与ConvNeXt-XL Cascade-Mask R-CNN相比, 在速度和精度方面分别提高了551%和0.7% AP, 在速度和准确率方面优于多种实时目标检测器。

Background:

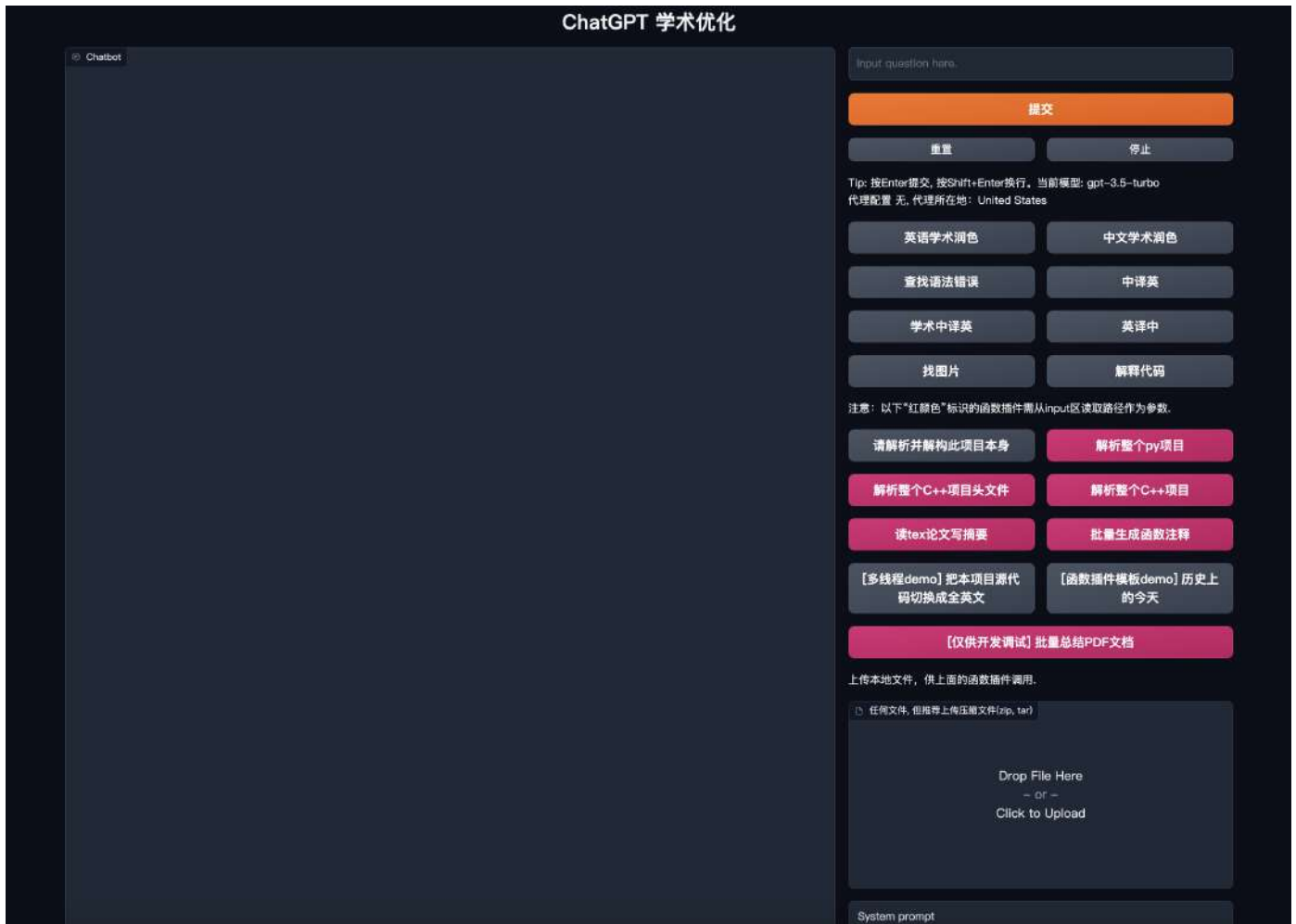
- a. Subject and characteristics:
  - 本文介绍了在计算机视觉领域中的实时目标检测方法的优化模型YOLOv7。
- b. Historical development
  - 在多个系统中, 实时目标检测是一个必要的组成部分, 它需要计算设备, 以执行从移动CPU和GPU到由主要制造商开发的神经网络单元 (NPU) 的实时目标检测。
- c. Past methods
  - 目前已有的实时目标检测算法有: YOLOv5, YOLOR, Scaled-YOLOv4, Deformable DETR, DINO-5scale-R50, ViT-Adapter-B等。
- d. Past research shortcomings
  - 现有方法存在精度和速度不足、计算代价过大等问题。
- e. Current issues to address
  - 目标检测算法中出现了新问题: re-parameterized模块替换初始模块和动态标签分配策略。

Methods:

- a. Study's theoretical basis
  - 本文的基础理论包括Cascade, group convolutions, re-parameterized模块等, 并使用优化模块和方法进行训练优化以提高目标检测的精度。
- b. Article's technical route (step by step)
  - 本文中提出的YOLOv7算法使用优化的卷积, Re-parameterized模块, 可靠的approximate non-maximum suppression方法, Dynamic Layer label Smoothing以及可训练的freebies方法等技术, 并融合了group convolutions, cascade结构, 道路轨迹以及深度监督等技术。

39、chatgpt\_academic (开源)

中科院学术版ChatGPT, 科研工作专用ChatGPT拓展, 特别优化学术Paper润色体验, 支持自定义快捷按钮, 支持markdown表格显示, Tex公式双显示, 代码显示功能完善, 新增本地Python工程剖析功能/自我剖析功能。

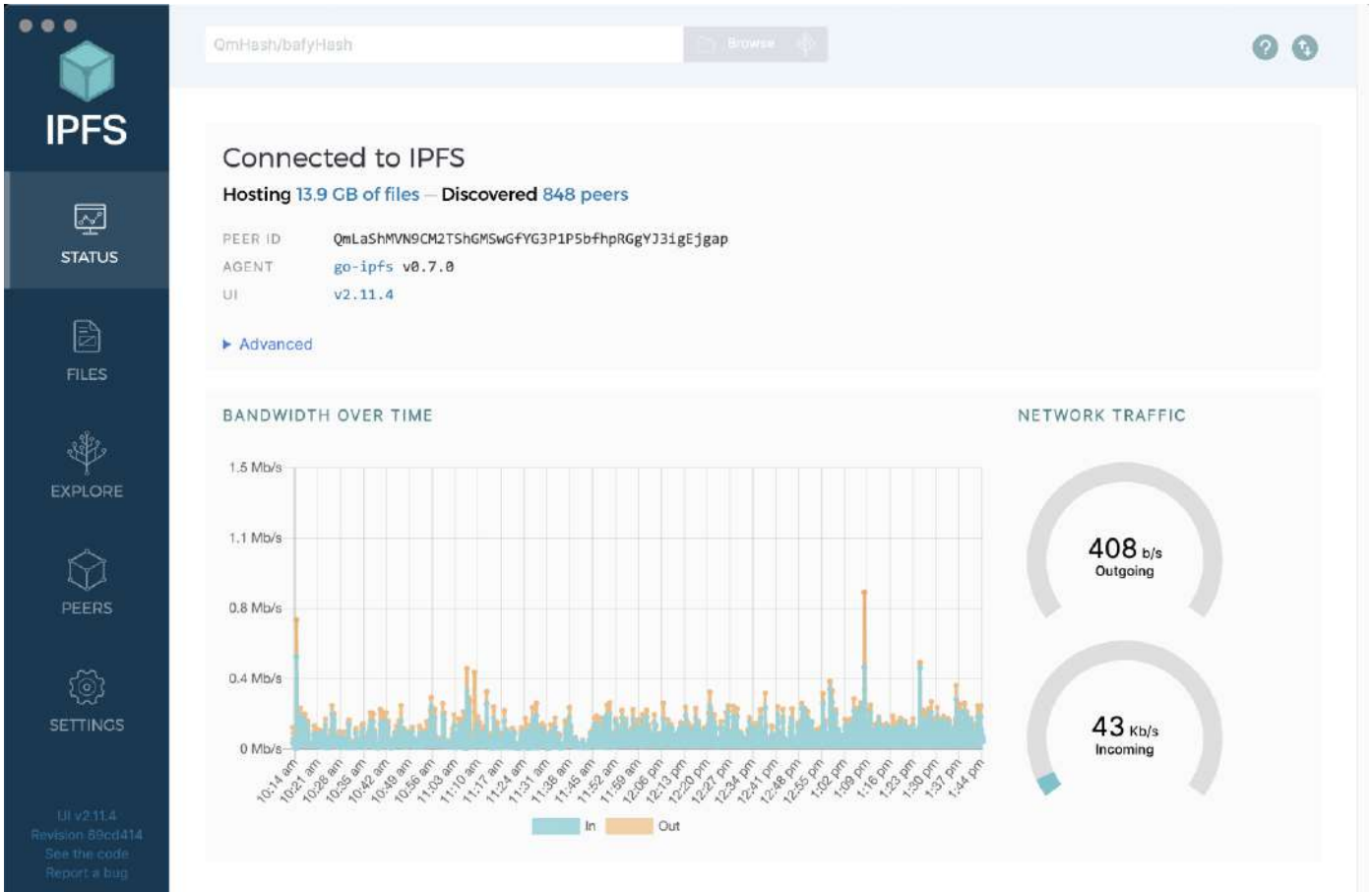


功能	描述
一键润色	支持一键润色、一键查找论文语法错误
一键中英互译	一键中英互译
一键代码解释	可以正确显示代码、解释代码
自定义快捷键	支持自定义快捷键
配置代理服务器	支持配置代理服务器
模块化设计	支持自定义高阶的实验性功能
自我程序剖析	[实验性功能] 一键读懂本项目的源代码
程序剖析	[实验性功能] 一键可以剖析其他Python/C++项目
读论文	[实验性功能] 一键解读latex论文全文并生成摘要
批量注释生成	[实验性功能] 一键批量生成函数注释
chat分析报告生成	[实验性功能] 运行后自动生成总结汇报
公式显示	可以同时显示公式的tex形式和渲染形式
图片显示	可以在markdown中显示图片
支持GPT输出的markdown表格	可以输出支持GPT的markdown表格
.....	

#### 40、[ipfs-desktop](#)

适用于 Windows、Mac 和 Linux 的 IPFS 的低调且用户友好的桌面应用程序。

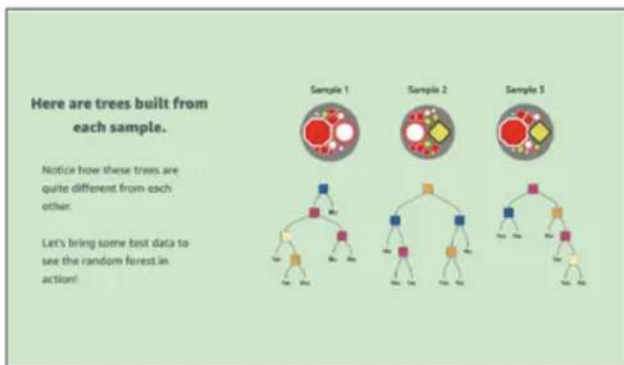




## 六、学习资源

### 1、机器学习概念的可视化解释（英文）

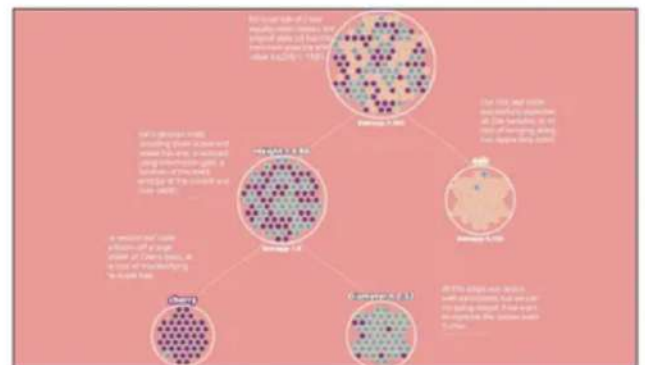
这个网站使用一系列短文，配上大量插图，解释机器学习的一些概念。



### **RANDOM FOREST**

Learn how the majority vote and well-placed randomness can extend the decision tree model to one of machine learning's most widely-used algorithms, the Random Forest.

[Dive In](#)



### **DECISION TREES**

Explore one of machine learning's most popular supervised algorithms: the Decision Tree. Learn how the tree makes its splits, the concepts of Entropy and Information Gain, and why going too deep is problematic.

[Dive In](#)

### 2、[SuperComputing-HPC-Data-Summary](#)（中文）

收录SC小组在学习高性能计算、分布式架构、数据挖掘与人工智能方向的笔记和材料。

### 3、[cdn-up-and-running](#) (英文)

从零开始构建 CDN 的教程。为了让你在实战中学习 CDN 的工作原理，这里会从创建一个单一的后端服务开始，逐渐扩展到多个节点、模拟延迟、可视化、可测试的 CDN 服务。因为设计 CDN 会涉及 Nginx、Lua、Docker、Grafana 等知识点，所以学习该教程需要有一定的编程基础。



### 4、[raft.github.io](#) (英文)

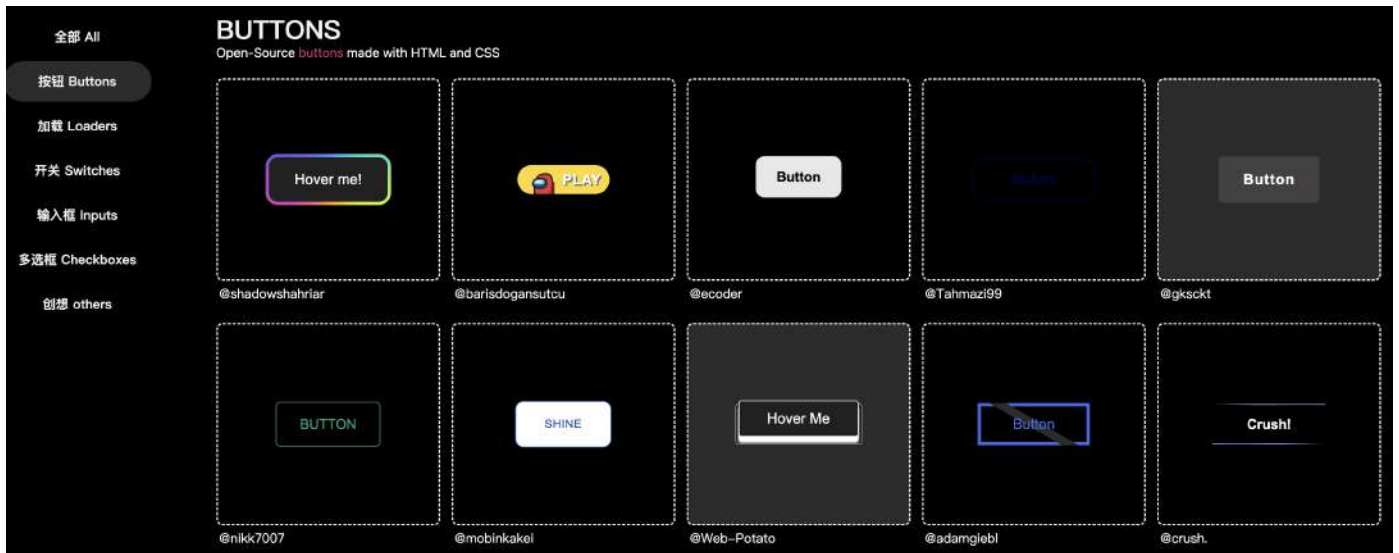
该网站收录了关于 Raft 的论文、课程、书籍等资料，以及相关开源项目和 Raft 的运行情况可视化，帮你彻底搞懂 Raft。

### 5、[algorithmica](#) (英文)

《现代硬件的算法》。该书来自俄罗斯非营利性的教育组织 Tinkoff Generation，它培养了大约一半的俄罗斯奥林匹克信息学决赛选手。不管你是算法研究员还是学生，这本书都可以让你学到更多提升程序性能的实用方法。

### 6、[handsome-css3-lib](#) (中文)

这个网站收集各种 CSS 特效和组件。项目地址：<https://github.com/ZiYi0414/handsome-css3-lib>



## 7、[Dash Dash](#) (中文)

这个网站将 Linux 文档 (man 文档) 重新格式化和排版，方便查找和阅读。

## 8、[GameDevMind](#) (中文)

全面的游戏开发技术图谱。该项目用思维导图的方式，展示了游戏开发需要具备的能力，包含技术栈、方法、工具、流程、管理、运营等方面。

## 9、[PPHC](#) (中文)

开源书籍《高并发的哲学原理》，这本书讨论的是 Web 服务高并发问题，内容由浅入深地介绍了 Apache、Nginx、epoll、交换机、k8s、数据库、分布式、微服务架构等解决高并发问题的技术和方案。

## 10、[k8s\\_PaaS](#) (中文)

教你用 K8s 部署一套完整服务的教程。通过该教程你可以学习到如何部署 Kubernetes 集群，以及在此基础上搭建由 Apollo、Jenkins、Prometheus 等服务组成的完整的软件研发和部署平台。



### 11、[小林 x 图解计算机基础](#) (中文)

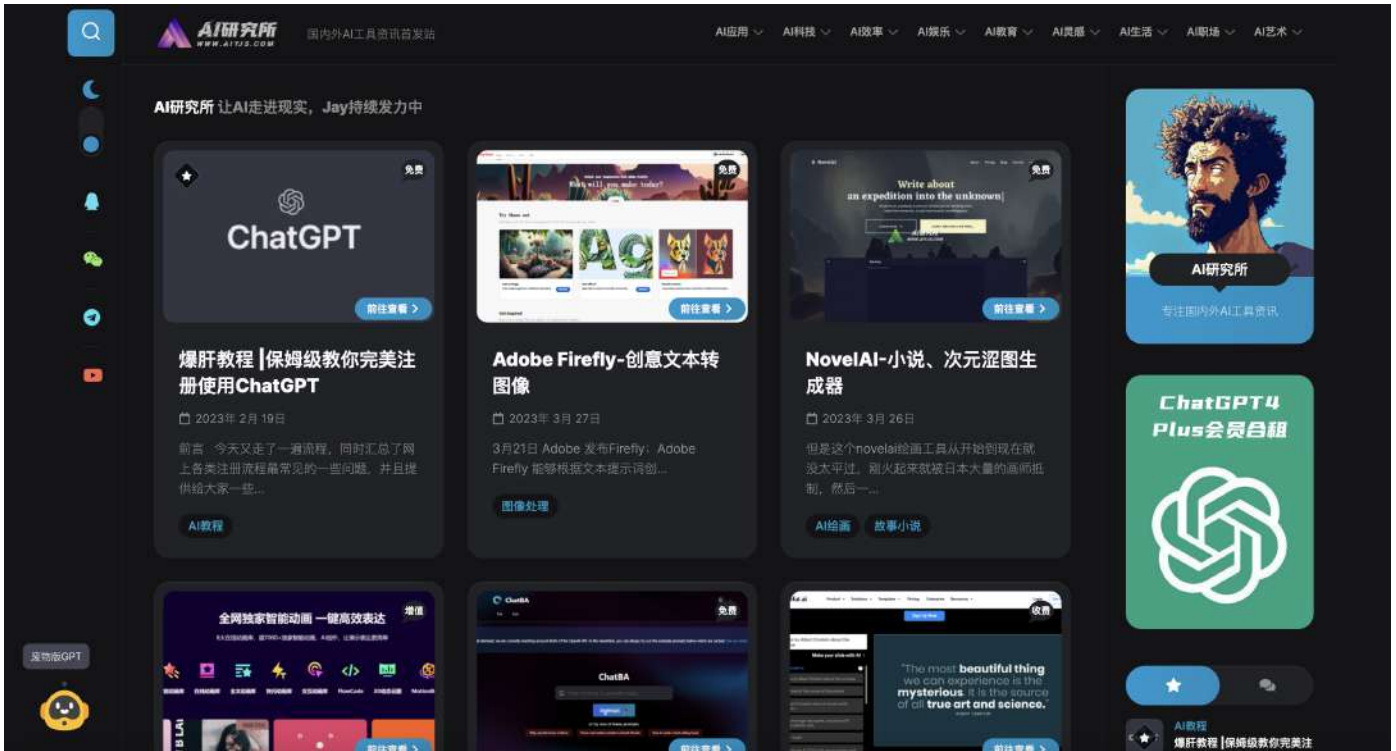
图解计算机网络、操作系统、计算机组成、数据库，共 1000 张图 + 50 万字，破除晦涩难懂的计算机基础知识，让天下没有难懂的八股文。

### 12、[Java全栈知识体系](#) (中文)

包含: Java 基础, Java 部分源码, JVM, Spring, Spring Boot, Spring Cloud, 数据库原理, MySQL, ElasticSearch, MongoDB, Docker, k8s, CI&CD, Linux, DevOps, 分布式, 中间件, 开发工具, Git, IDE, 源码阅读, 读书笔记, 开源项目...

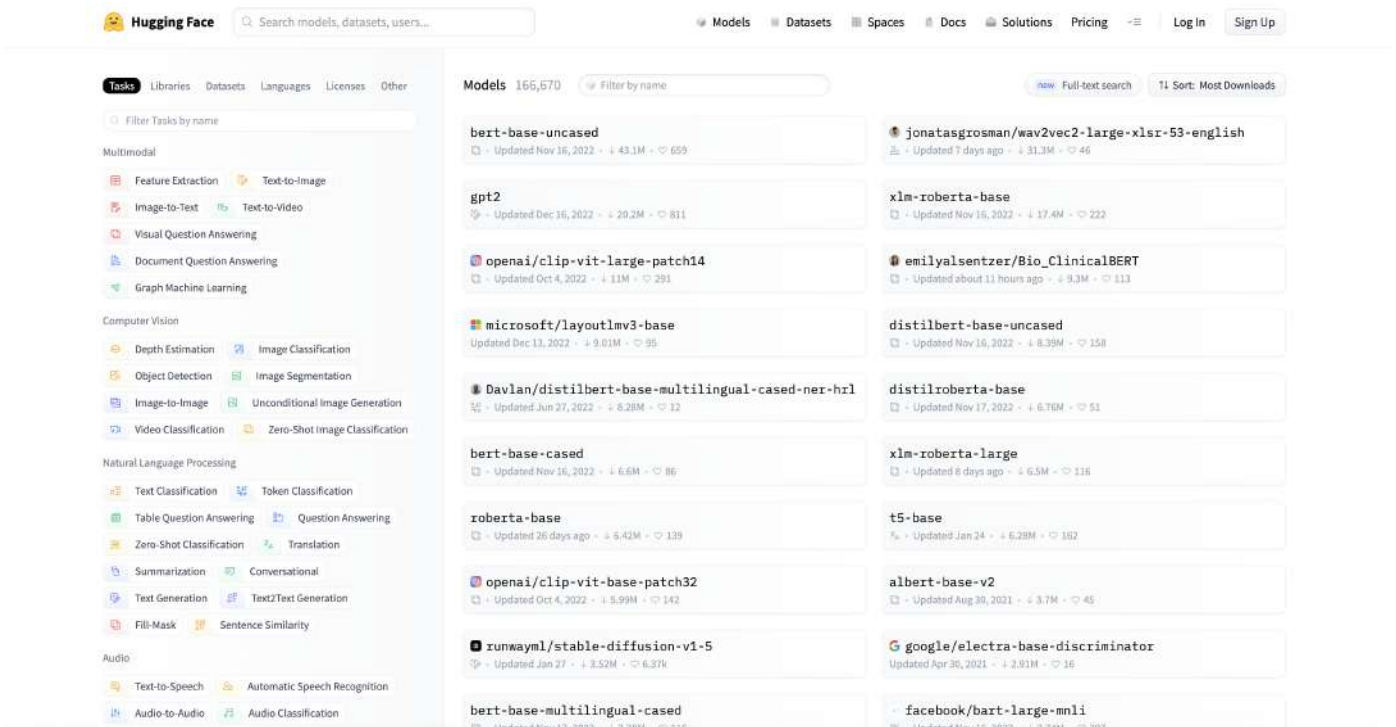
### 13、[AI研究所](#) (中文)

一个中文网站，收录 AI 相关工具。



## 14、Hugging Face (英文)

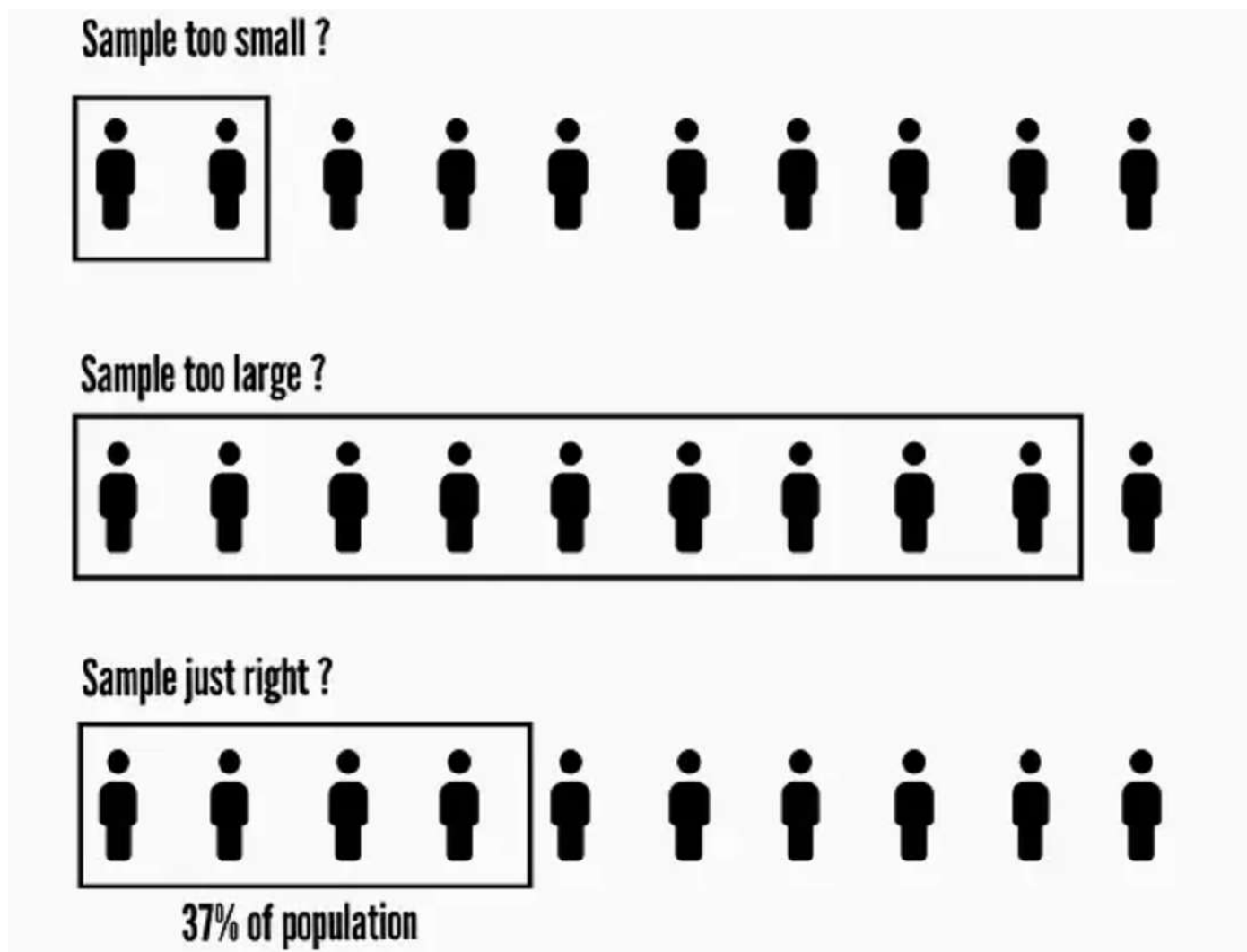
一个深度学习领域的社区，主要用于寻找数据集及预训练模型。Transformers 提供了数以千计的预训练模型，支持 100 多种语言的文本分类、信息抽取、问答、摘要、翻译、文本生成。它的宗旨是让最先进的 NLP 技术人人易用。



## 七、生活经验

### 1、“1/e 法则”

某公司招聘一名秘书，有100名候选人，依次面试。每面试完一个人，就必须立刻决定是否录取。也就是说，不能面试完所有人，再回过头决定录取哪一个，一旦放弃当前候选人，就只有从后面的面试者中选择。这个设定是合理的，象征我们在生活中遇到的各种机会。机会来临时，转瞬即逝，必须立刻决定是否抓住它，错过就是错过了。如果录用得太早，可能错过后面更好的候选人；如果录用得太晚，可能错误放走前面的合格人选。



数学家对这个问题，已经有了充分的讨论。经过计算，成功概率最大的方法，叫做“ $1/e$  法则”，它约等于37%，所以又称“37%法则”。它的意思就是，寻找阶段进行到37%就要停止。100个应聘者，先面试前37个，此后的面试只要遇到一个更优秀的，就立刻录取，不再继续面试了。换句话说，前37个面试者无论多么优秀，都不会录取，他们只是用来确定录取的标准。

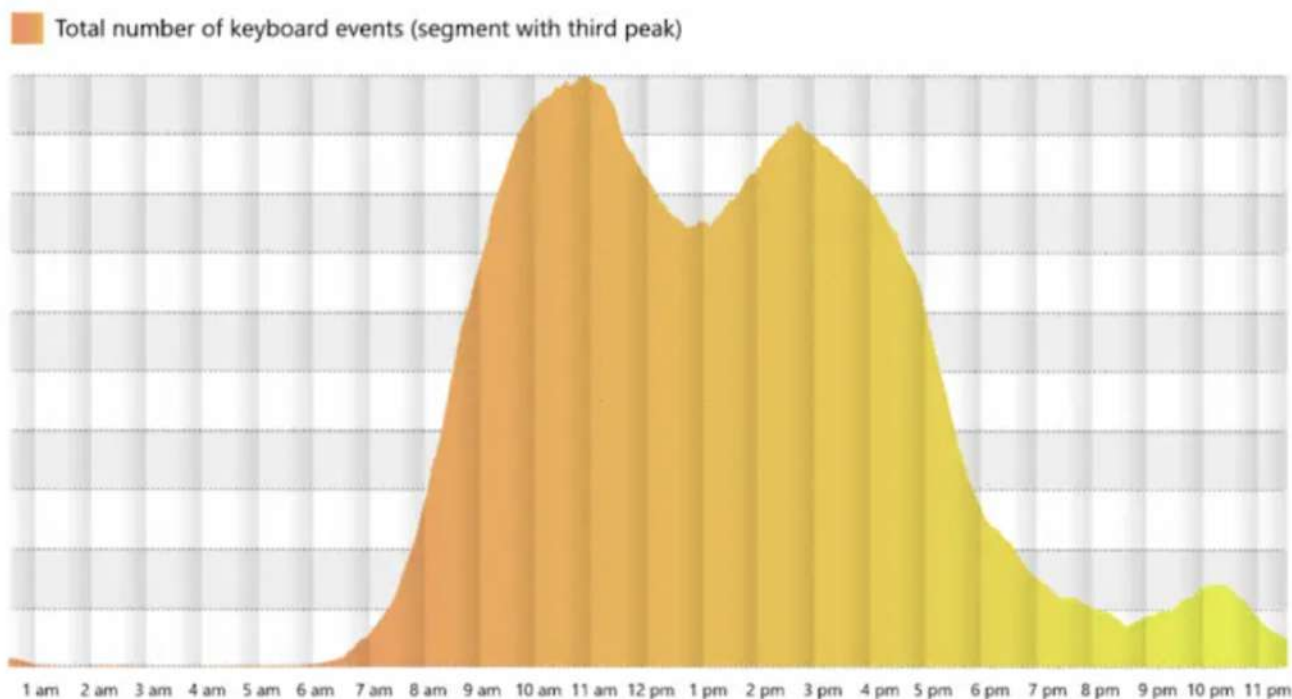
如果最合适的候选者偏偏在前面37%里面，那就只能错过了，作为“寻找阶段”不得不付出的成本。最终录取的将是不如前面候选人的次优选择。这个法则很实用，日常生活中，只要符合“寻找-决策过程”的场景，都可以适用37%法则。

- (1) 相亲时，假定有10个相亲对象，那么前3~4个可以作为寻找阶段，后面只要遇到一个比前面更好的人，就可以同意了。
- (2) 租房时，假定有一个月的找房子时间，那么30天的37%也就是11天。在找了11天之后，你就要出手了。只要发现比先前更令人心动的房子，就不要犹豫，马上租下来。
- (3) 读书时，假定这本书有100页，如果读了37页，还没有发现感兴趣的内容，那就可以放弃了。
- (4) 一个10集的电视剧，第4集是最佳弃剧时间。
- (5) 一个10分钟的视频，看了3分42秒，如果还是觉得不好看，就可以关掉了。

(6) 一个年轻人想在18岁到24岁，一共7年时间里找到人生方向，确定未来想做什么。那么，他有2.59年（ $7 * 0.37$ ）的时间自由尝试。也就是说，到了大三下学期就应该初步定下自己的方向，后面除非遇到更有吸引力的事情，否则就不应该转换事业方向。

## 2、居家办公的3个高峰时段

微软公司的 Office365 团队，统计了各家公司远程网上办公的数据，发现在家办公每天有三个高峰时段。第一个高峰是上午10点到11点，第二个是下午2点到3点，第三个是晚上10点到11点。



## 3、世界前100万名网站

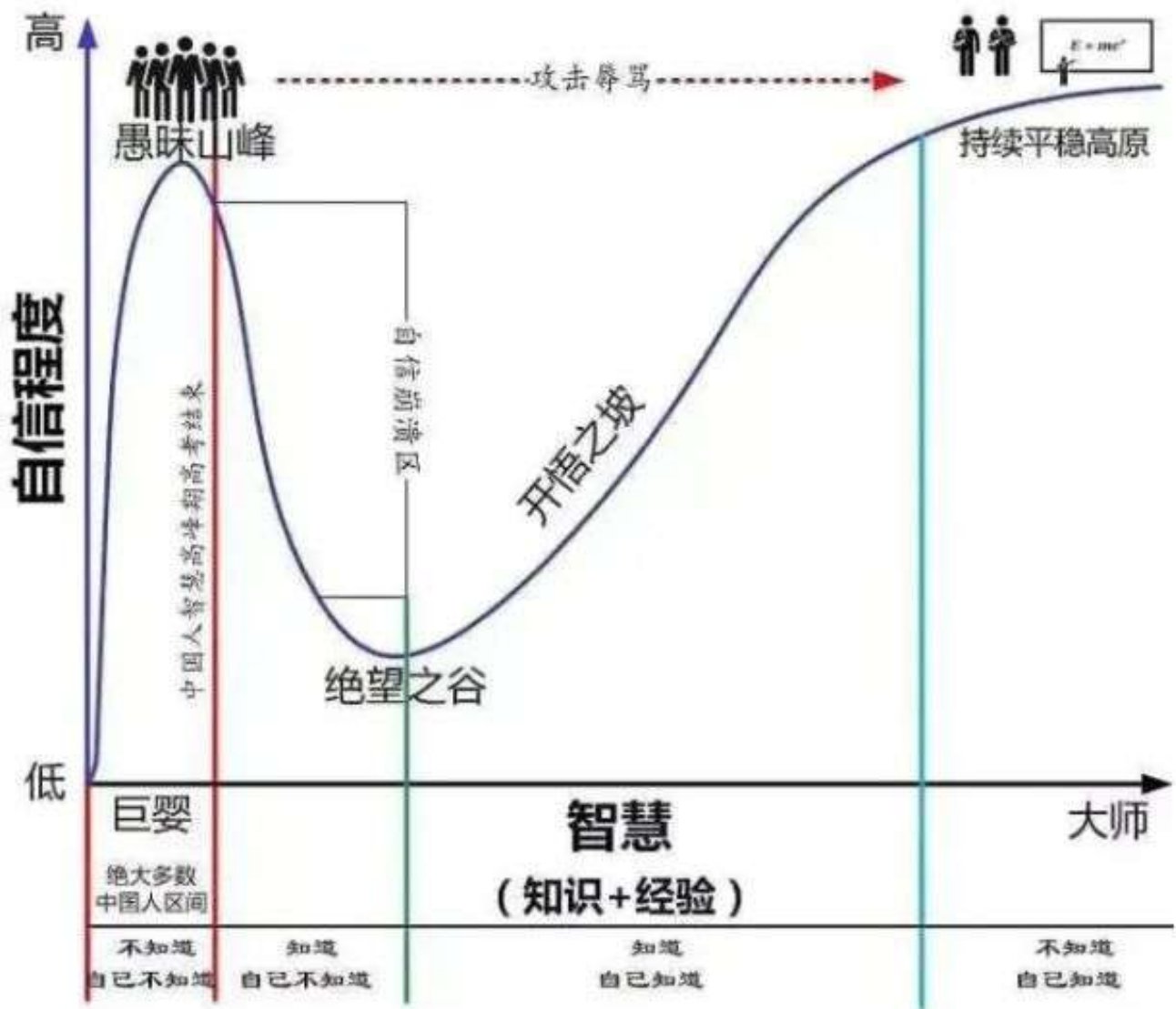
谷歌公司根据 Chrome 浏览器的数据，公开了世界前100万名网站排名，链接里的仓库提供这个名单的下载。不过，因为 Chrome 浏览器在中国受限，它的中文网站排名并不准确。

## 4、达克效应

达克效应（DK effect）是一种认知偏差，能力欠缺的人有一种虚幻的自我优越感，错误地认为自己比真实情况更加优秀。美国康乃尔大学的社会心理学家大卫·邓宁和贾斯汀·克鲁格将其归咎于元认知上的缺陷，能力欠缺的人无法认识到自身的无能，不能准确评估自身的能力。他们的研究还表明，反之，非常能干的人会低估自己的能力，错误地假定他们自己能够很容易完成的任务，别人也能够很容易地完成。但能力欠佳的人对自己能力的评估并不比能力较佳的人高。邓宁和克鲁格于1999年在实验中首次观测到此认识偏差，他们通过对人们阅读、驾驶、下棋或打网球等各种技能的研究发现：

- 能力差的人通常会高估自己的技能水准；
- 能力差的人不能正确认识到其他真正有此技能的人的水准；
- 能力差的人无法认知且正视自身的不足，及其不足之极端程度；
- 如果能力差的人能够经过恰当训练大幅度提高能力水准，他们最终会认知到且能承认他们之前的无能程度。

邓宁和克鲁格认为这种效应是由于能力欠缺者的内在错觉和能干者对外界的错误认知：“无能者的错误标度源自于对自我的错误认知，而极有才能者的错误标度源自于对他人的错误认知。”



## 5、[土耳其的地下城市](#)

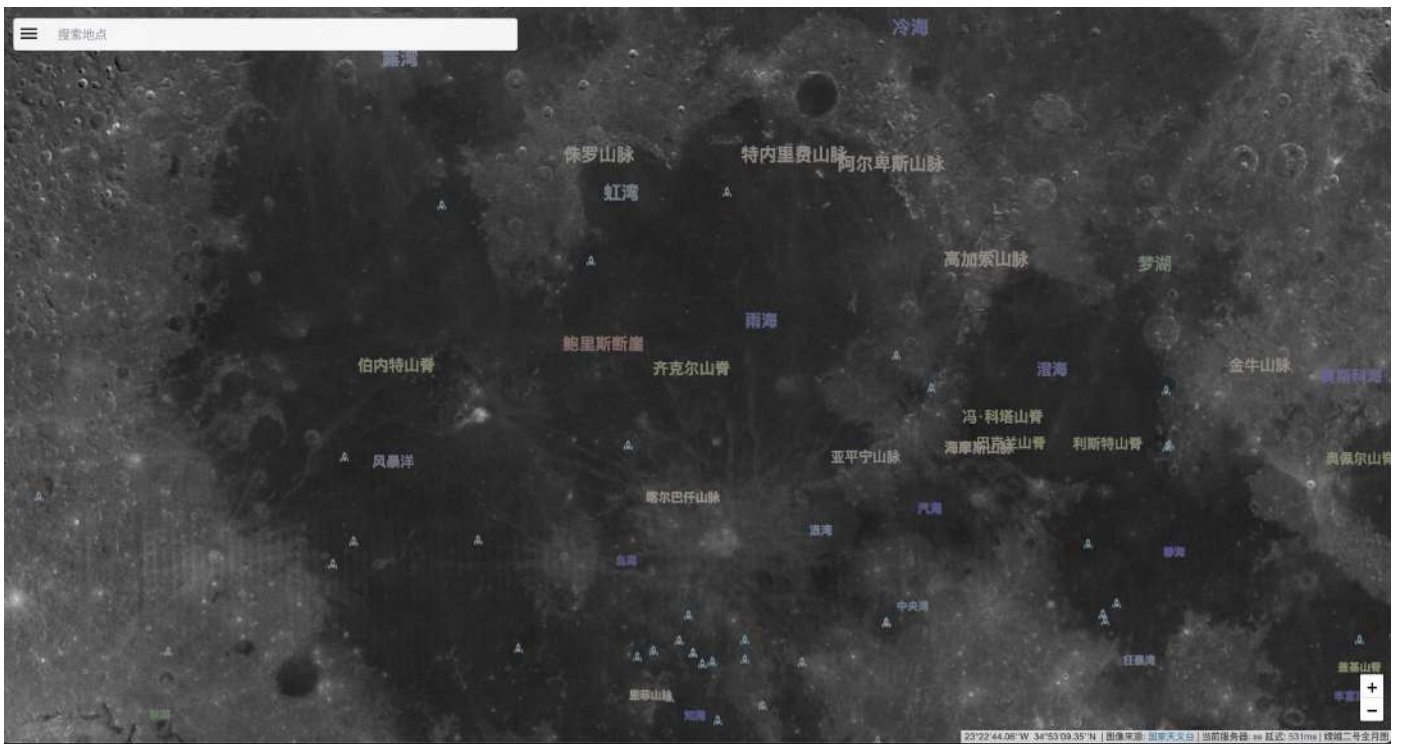
1963年，土耳其德林库尤镇的一个男子装修地下室时，意外发现地下室的下方有隧道，从而揭开了一个庞大的地下城市。这座地下城市有18层，最深处达到76米，可以容纳两万人。一些历史学家认为，这座城市是公元6世纪至10世纪时，当地人为了逃避宗教迫害而建。另一些历史学家则认为，这个镇所在的卡帕多西亚地区，冬天非常寒冷，夏天非常炎热，居民选择住在地下，这样就温度适中。





## 6、[月球地图](#)

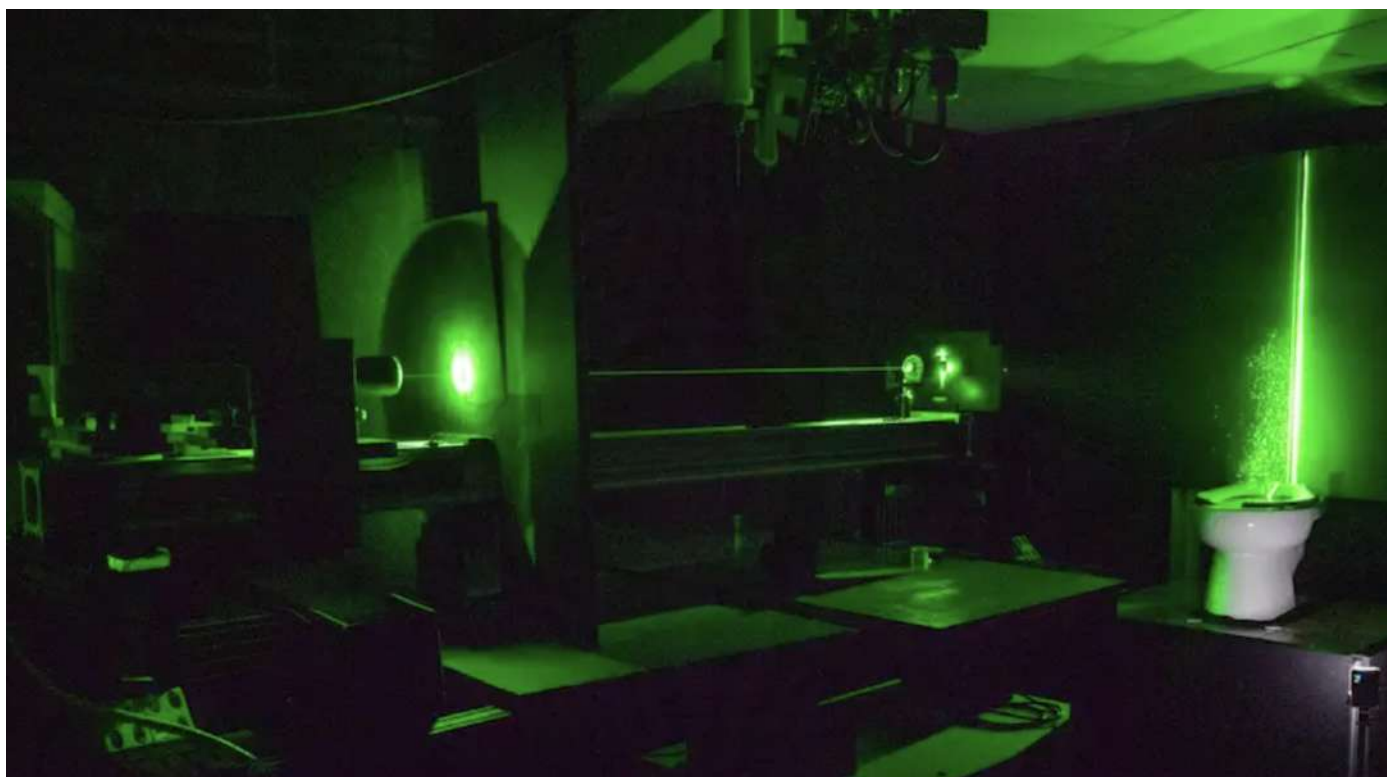
使用嫦娥二号发布的全月影像图制作，分辨率可以达到7米，能够搜索和定位各种环形山、地名以及人造物体位置。



## 7、[马桶冲水的气溶胶](#)

公共卫生学家一直呼吁，马桶冲水会产生气溶胶的粒子扩散，可能会传播病原体，但是没有证据证实这种扩散的速度和分布。

科罗拉多大学的工程师团队，使用绿色激光和高速摄像设备，拍摄马桶冲水，证实确实存在气溶胶扩散。实验发现，冲水时，颗粒以每秒2米的速度快速射出，可以到达马桶上方1.5米的高度。虽然较大的颗粒会在几秒钟内沉降到马桶表面上，但较小的气溶胶颗粒可以在空气中悬浮几分钟或更长时间。下图是实验时的激光拍摄场景。



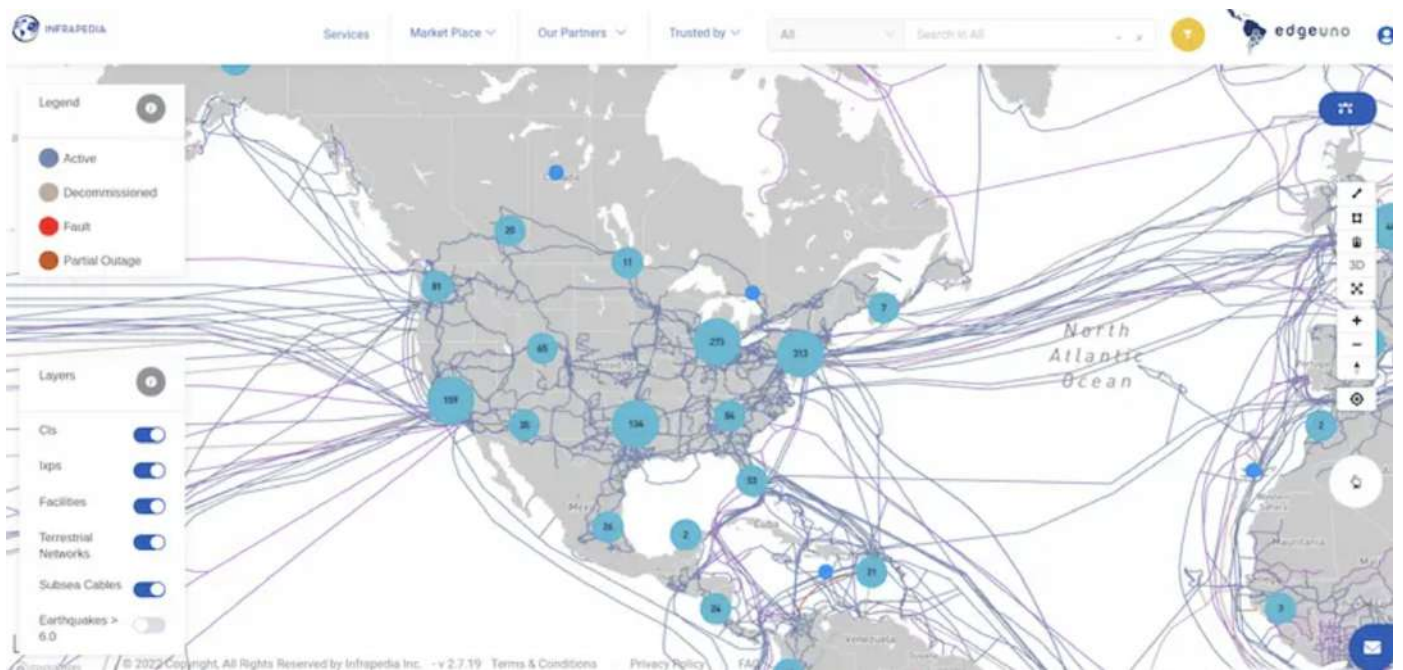
## 8、[GPS 艺术](#)

谷歌地图会根据 GPS 信号，显示运动轨迹。很多人就通过这种方式，在谷歌地图作画，称为 GPS 艺术。一个日本男子为了向女友求婚，在谷歌的日本地图上写下了“Marry Me”，最后还画了一颗被箭射中的心。他从日本北部的北海道，一直开车到南端的鹿儿岛，一共行进了7163公里，为期6个月，是世界最大的 GPS 艺术品。



## 9、[世界互联网基础设施地图](#)

这个网站可以查看全世界的主干光纤线路和机房位置。



## 八、闲情逸趣

### [大数据已死](#)

谷歌的大数据工程师乔丹·蒂加尼（Jordan Tigani），最近直言不讳地说：“大数据已死”。他认为，大数据时代已经结束了，大数据的存储和分析，作为一个技术问题已经解决了。用户已经不必担心数据大小了，再多的数据都不是问题。

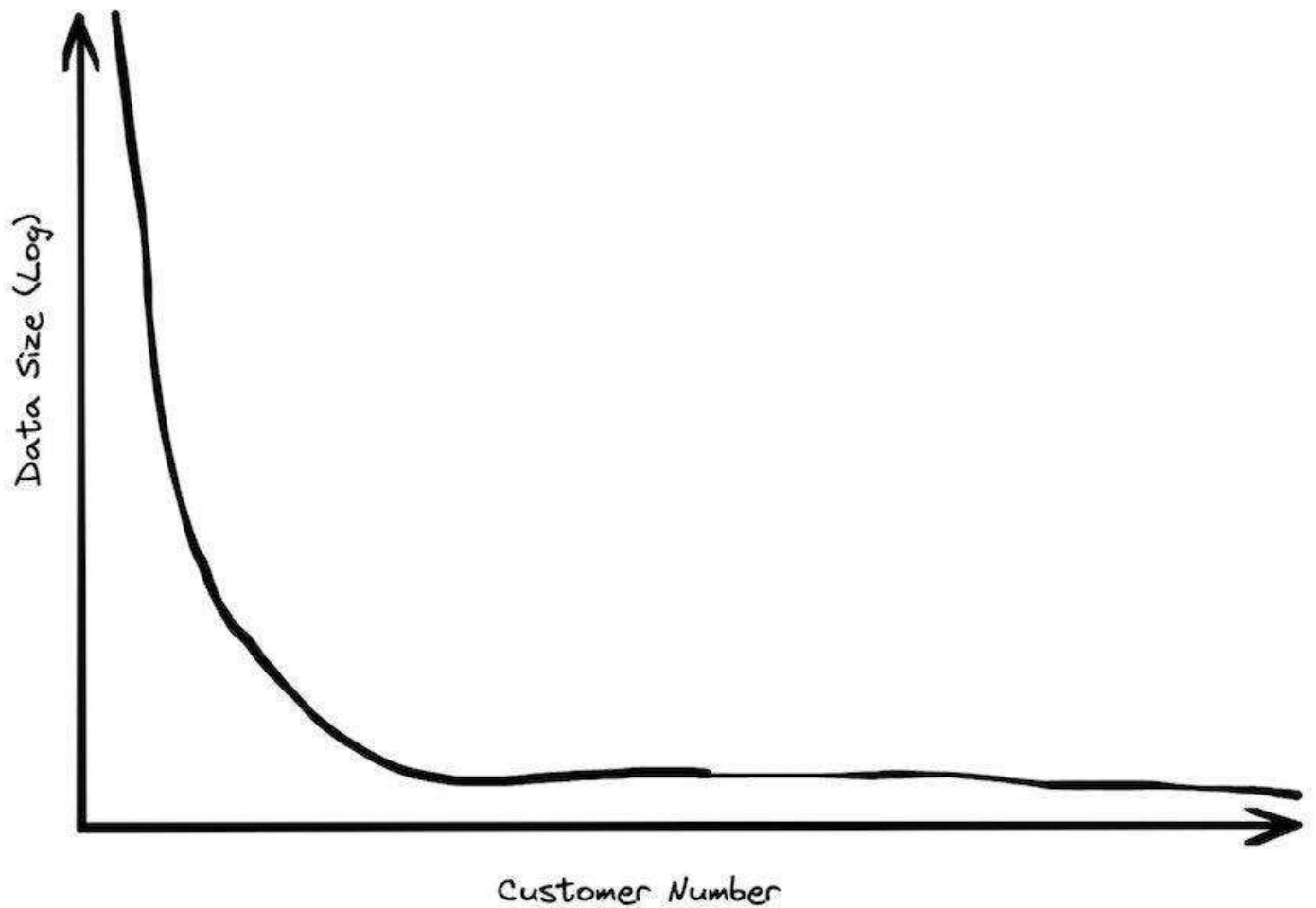


下面是他提出“大数据已死”的6个理由。

[1] 绝大多数企业到不了大数据级别。企业的数据量往往不到 1TB，很多甚至不到 100GB。

假设一家中等规模的制造业公司，拥有1000个客户，每个客户每天产生一个订单，每个订单包含100个产品。这家公司一天产生的数据量，依然远远小于 1 MB。三年后，数据总量也只有 1 GB，达到 1 TB 需要几千年。

就算是大型互联网公司，大多数时候也到不了大数据级别。假设某个营销活动有100万用户参加，并且同一时间，该公司开展了几十个这样的营销活动，每天的数据量依然不足 1 GB，就算加上各种日志，可能也只有几个 GB，这跟大数据相差甚远。



[2] 存储和计算正在分离。大数据包含“数据存储”和“数据计算”两个方面，如果放在一个系统里面处理，确实很难。但是，这两方面现在已经能够脱钩，变成两个独立系统，各自都能独立扩展。这意味着，“数据计算”不受“数据存储”（数据库大小）的限制，反之亦然。因此，大数据作为单一问题就不存在了，变成了海量存储和大型计算两个问题。

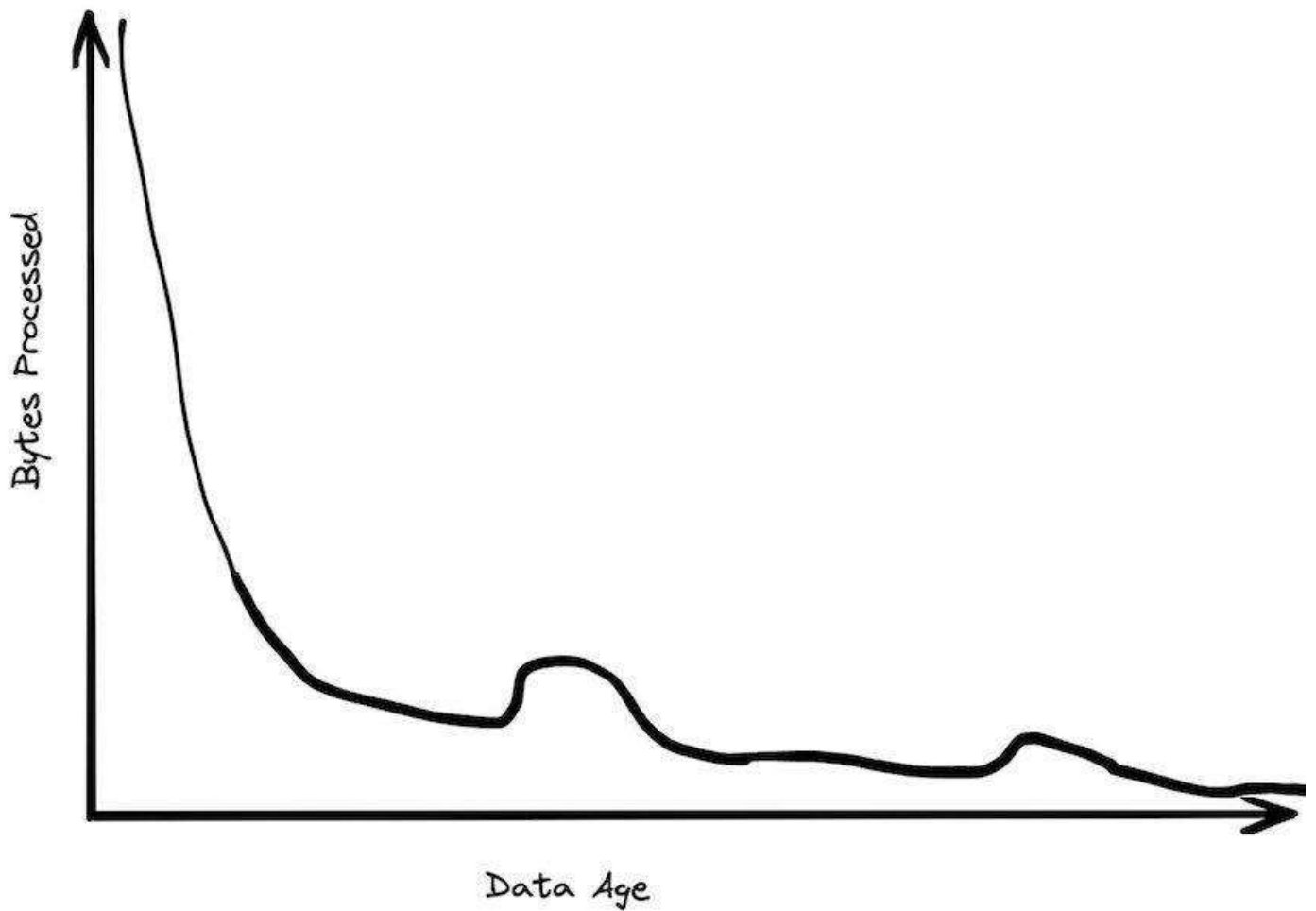
[3] 没有新业务的情况下，数据是线性增长的，即每天的新增数据与以前的数据结构相同。

以前的数据一旦写入数据库，通常就不再发生变化，也没有新的计算需求，相关计算在以前就完成了。这时只要对最近的新增数据进行单独计算，然后保存就可以了。你很少需要每天扫描一遍旧数据：那些数据一成不变，为什么要一遍一遍计算它们呢？

因此，对于一家企业来说，“数据会指数式增长”这个假设并不成立。而且，数据计算的需求，其实比数据存储的需求小得多，因为老数据很少需要再次计算。

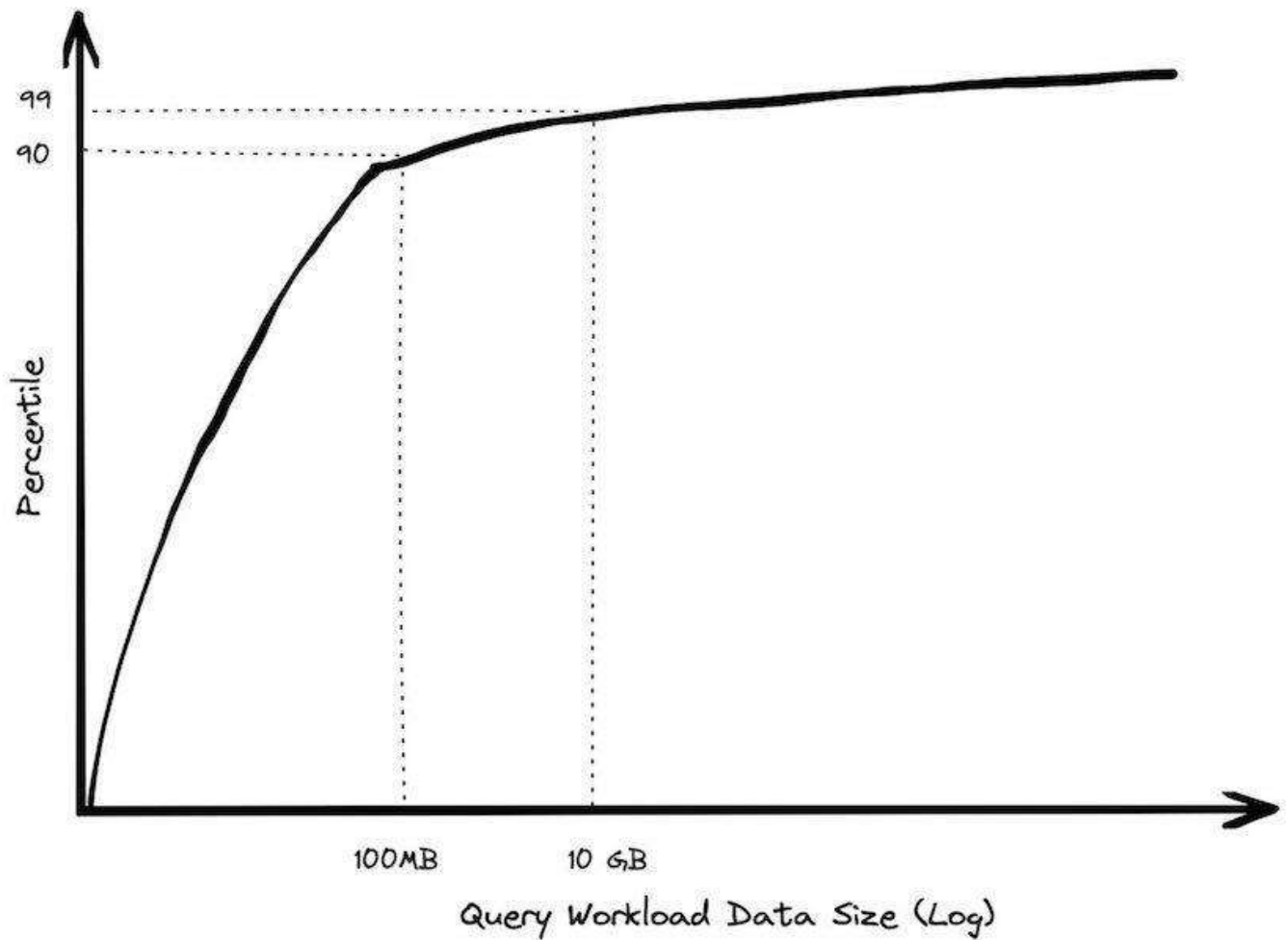
[4] 人们看重的往往只是最近的数据。最频繁的查询是针对24小时内产生的数据，一周前的数据的查询可能性要低20倍，一个月前的历史数据只会偶尔被查询。

这意味着，大数据更像静态数据，而不完全是动态数据。既然以前的数据很少用到，那么就可以压缩保存。一个包含10年数据的表格，可能会达到PB级别，但是如果压缩保存历史数据，压缩后可能不到50GB。



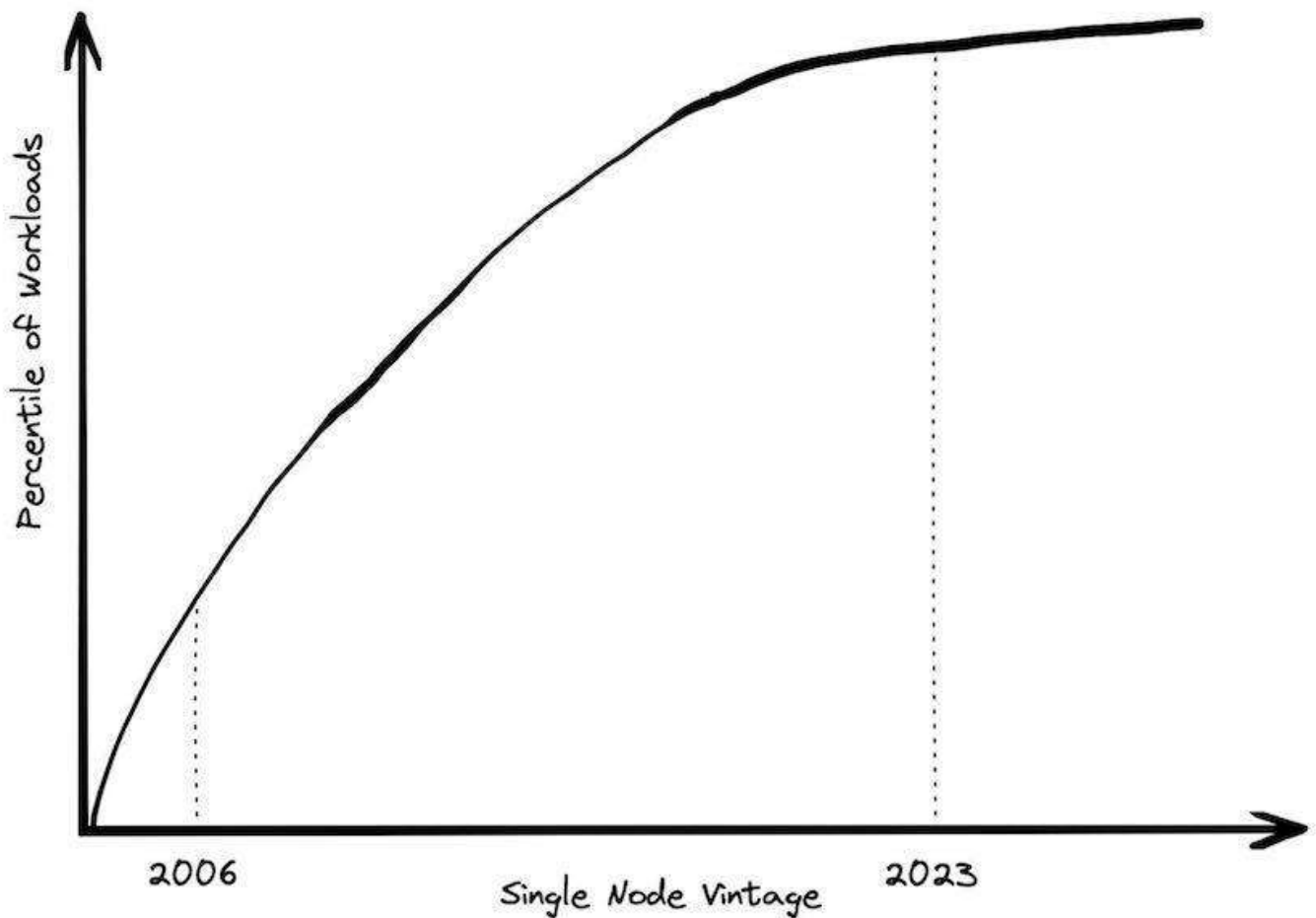
[5] 真正拥有大数据的公司，几乎从不查询全部数据。他们90%的查询涉及的数据少于 100 MB，涉及 TB 级别数据的查询非常少。

就算查询 TB 级别数据，查询性能的优先级往往并不高。等一个周末或几天才拿到结果，通常是可以接受的。另外，大型数据集的查询非常昂贵。谷歌的 BigQuery 的 PB 级别查询报价是 5,000 美元，即使是大公司也不会经常使用。



[6] 硬件的飞速发展，使得单台计算机的计算能力大增。2004年，谷歌发表 MapReduce 论文时，单机的计算能力还比较弱，很多计算必须通过分布式完成。

2006年，AWS 推出了 EC2 云主机，你只能用到一个单核 CPU 和 2 GB 内存。今天，AWS 的标准实例具有64个内核和 256 GB 内存。如果愿意多花钱，还可以拿到445个内核和超过 24 TB 内存。单机计算能力大大增强，意味着大数据的最大难点分布式计算，即使被用到，困难程度也大大降低。



综上所述，结论就是：数据量已经不需要特别关注了，再也不必担心处理不了海量数据了。大数据作为一个技术问题，已经解决了。

## 九、数字与言论

1、拍电影时，最重要的是你知道你想用电影表达什么。很多导演追求使用高科技拍电影，但对于他们真正想要表达的内容，却只有一个非常模糊的概念。人们其实非常欣赏那些制作技术并不先进、但知道自己想说什么的电影。——[宫崎骏](#)

2、大多数编程不再是艺术，也不再是高级工程。大多数程序都是微不足道的：编码人员只需要理解某些特定的接口，然后编写一些逻辑和胶水代码。——[Antirez](#)，Redis 的作者

3、领导者的作用不是行使权力，而是分配权力。有效的领导者通过设定方向、确定人员、授权他人、确保决策执行，来节省自己的时间，并扩大自己的影响。——[《我的创业箴言》](#)

4、开发 Chrome 浏览器时，有一次我们讨论，应该允许网页重定向多少次。有人提出设为10次，超过这个数字，Chrome 就会放弃加载网页。然后，另一个人说不行，必须设为30次，否则《纽约时报》网站就会停止工作。于是，Chrome 的重定向上限就是30次。——[《我在 Chrome 团队的10年》](#)

5、摩尔定律只说了，每18个月计算机芯片的晶体管数量增加一倍，但是没有说增加一倍所需要投入的科研人员，今天是1970年代的18倍。——[《2022 年生物学突破》](#)

6、生命太短暂，不能花在那些不值得阅读的内容上面。就算你是一个很爱读书的人，活到70岁最多大概能阅读15000本书，这只占世界最大图书馆美国国会图书馆3800万册藏书的0.04%。我们一生中能够阅读的书籍其实很少。因此，关键技能不是多读，而是跳过那些不值得读的内容。——[Hacker News](#)



7、有些领域变化非常快，在有人写书之前，博客有时是唯一的信息来源。Stable diffusion 模型出现后的第二天，人们就已经在写博客了，书籍永远不会那么快。而且，博客往往是免费的，而书籍和论文则被锁定在付费墙之后。因此，你可以这么认为，博客获取灵感，书籍获取知识。——[Hacker News](#)

8、比特币的原始设想，只是提出了一个有趣的去中心化实验，但是等到有人发现可以从中获利，事情就变质了。这原本只是一群程序员在玩加密游戏，后来有人创建了纯粹的营利性企业，比特币从此就变了。——[《加密货币与皇帝的新衣》](#)

9、衡量一个人的领导能力的最好方法，就是看如果这个人休假了，他的下属在做什么。优秀的产品经理和工程师可以休假一周，他管理的工作不发生任何问题。优秀的主管和技术负责人可以休假一个月。领导能力越优秀，休假的时间就越长。——[Andrew Bosworth](#)，Facebook 的 CEO

10、简洁的写作需要勇气。让事物变小是一种深思熟虑的、困难的和有价值的行为。——[《发布 Newsletter 35年的35条教训》](#)