

GPT4ALL 启动会

王伟

X-lab & OpenEduTech

2024 年 1 月



背景：全民数字素养与技能提升 → 数字化学习平台

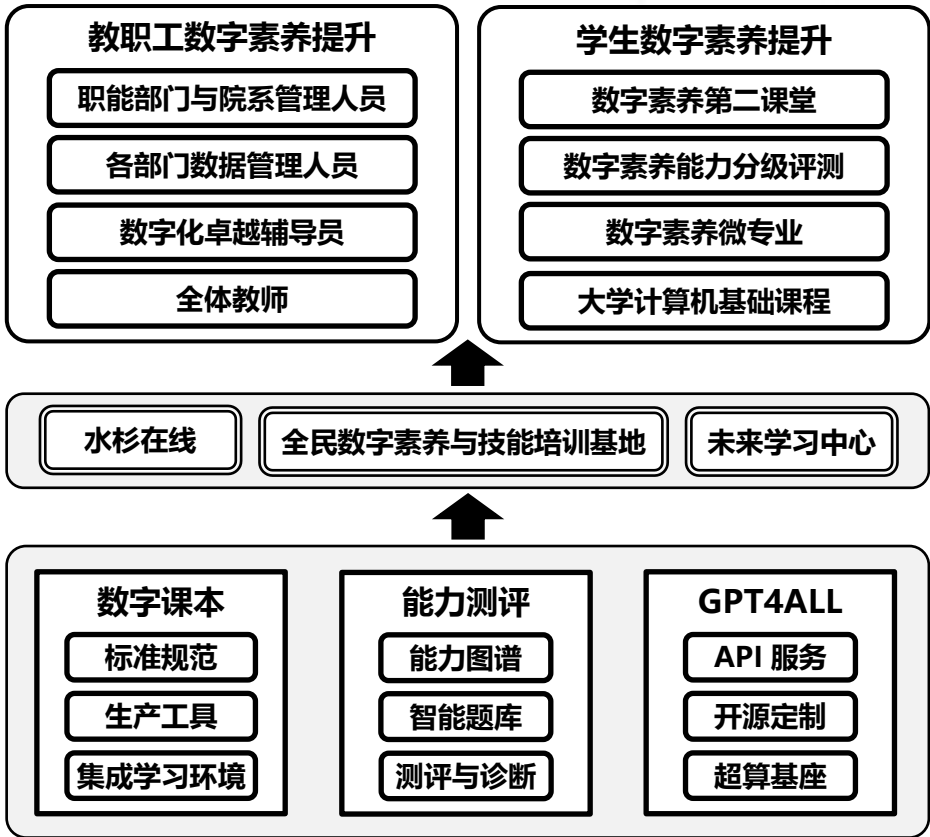
教-学-练-测-创-评

全民社会服务

跃升服务

平台与数据

关键技术



面向“全民数字素养提升”的教育大模型



教育场景

AI4Coding

AI4SE

AI4DS

AI4OSS

代码级生成

任务级生成

场景级生成

关键技术

测评与优化

测评规范

测评工具

优化方法

边缘与轻量

AIGC一体机

适配工具链

BigDL 项目

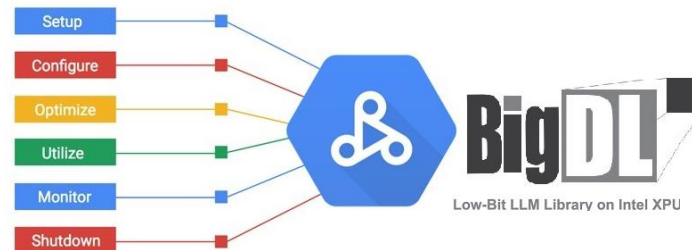
GPT4ALL

教育场景适配

提示词工程

低代码平台

基础设施



GPT4ALL 研究框架



教育场景 AI 能力域	教	学	练	测	创	评
基础知识 (AI4DL)	<ul style="list-style-type: none"> 教案生成 出题组卷 知识图谱 	<ul style="list-style-type: none"> 自适应学习 智能助教 	<ul style="list-style-type: none"> 个性化练习 自动纠错 拓展练习 	<ul style="list-style-type: none"> 自动批改 个性化试题 自动打标签 	<ul style="list-style-type: none"> 创意生成 项目推荐 智能查新 	<ul style="list-style-type: none"> 测评大模型 能力画像
编程技能 (AI4Coding)	<ul style="list-style-type: none"> 编程知识图谱 	<ul style="list-style-type: none"> 代码助教 	<ul style="list-style-type: none"> 代码纠错 	<ul style="list-style-type: none"> 代码自动出题 代码自动打标 	—	<ul style="list-style-type: none"> 代码能力评测
工程能力 (AI4SE)	<ul style="list-style-type: none"> 工程知识图谱 	<ul style="list-style-type: none"> DevGPT 	<ul style="list-style-type: none"> AIOps 助教 	<ul style="list-style-type: none"> DevOps 评测 	<ul style="list-style-type: none"> 工程场景生成 工程管理 Agent 	<ul style="list-style-type: none"> 工程能力评测
数据分析 (AI4DS)	<ul style="list-style-type: none"> 数据科学图谱 	<ul style="list-style-type: none"> GPT-DB GPT-GDS 	<ul style="list-style-type: none"> 实训数据生成 分析代码生成 	<ul style="list-style-type: none"> 分析模型评测 SQL/GQL生成 	<ul style="list-style-type: none"> 数据作品辅助 数据场景生成 	<ul style="list-style-type: none"> SQL/GQL评测
开源协作 (AI4OSS)	<ul style="list-style-type: none"> 开源知识图谱 开源项目标签 	—	<ul style="list-style-type: none"> GitOps助教 	<ul style="list-style-type: none"> 开源社区问答 	<ul style="list-style-type: none"> 开源协作机器人 	<ul style="list-style-type: none"> 贡献度评价 开源人才画像
基础设施 (AI4Infra)	<ul style="list-style-type: none"> 基础代码大模型、代码大模型评测、超轻量大模型 (BigDL)、教育场景低代码平台、大模型一体机 					



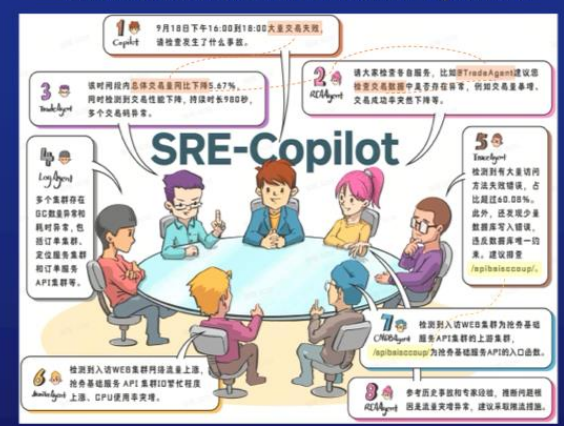
SRE-Copilot架构简介

SRE-Copilot是基于LLM的多场景智能运维框架, 支持Multi-Agent协作与动态编排, 具备计划、记忆、反思、推理与ReAct等能力, 为SRE提供智能化服务。



主要运维能力: 异常检测

——基于ReAct框架与思维链CoT的Multi-Agent编排



角色定义

- Copilot主持人:** 解析用户需求, 制定运维Plan, 安排不同Agent工作 (如根据定位交给RCA)
- 多数据源Agent:** 分别负责不同模态的数据, 选择合适的算法进行异常检测与检索
- RCAAgent:** 收集其他Agent检测到的异常信息与链路、配置信息, 进行根因定位

Multi-Agent编排

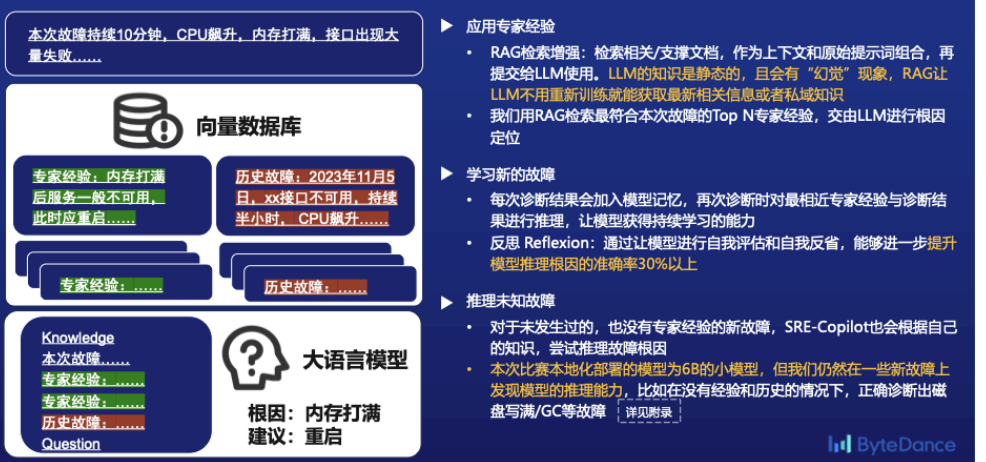
- ReAct包括推理Reasoning和行动Action, 推理帮助模型生成、追踪和更新计划并处理异常, 行动允许模型与外部环境交互以获取更多Observation, 提升准确率与适应性
- 每个Agent均根据检测到的异常动态编排, RCAAgent负责收敛协作轮次, 并根据其他Agent反馈决定下一步分析与下钻的方向

优势

- 模拟真实云平台跨组件协同定位, 可以采用多个agent替代多个组件运维团队
- 降低噪音, 鲁棒性更高: 异常检测关注异常时候多个组件的表现形态, 而非根据单一组件判断是否异常

主要运维能力: 根因定位

——基于RAG增强的推理与反思



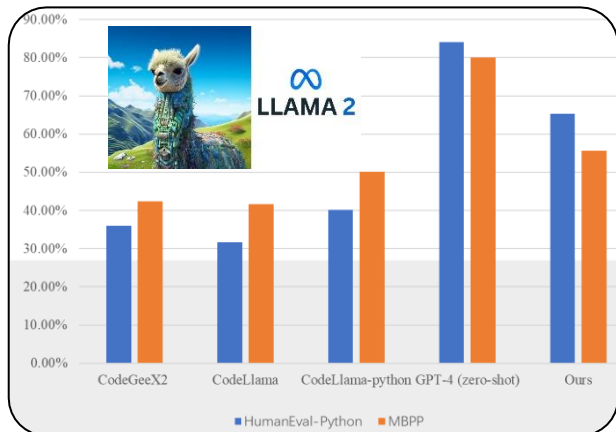
Birmingham 2023
CIKM
The 32nd ACM International Conference on Information and Knowledge Management
October 21-25, 2023
Birmingham, UK

Best Demo Paper Award

Jianing Wang, Nuo Chen, Qiushi Sun, Wenkang Huang, Chengyu Wang & Ming Gao
HugNLP: A Unified and Comprehensive Library for Natural Language Processing

The General Chairs: *Mark Lewis*
The Demo Paper Program Chairs: *Ming Gao*

水杉代码大模型



大模型实训一体机

DL4ALL 全民数字素养与技能培训基地

给定一个整数数组 `nums` 和一个整数目标值 `target`，请在该数组中找出和为目标值 `target` 的那两个整数，并返回它们的数组下标。你可以假设每种输入只会对应一个答案。但是，数组中同一个元素在答案里不能重复出现。

输出: [2, 7]

复杂代码生成

```
class Solution:
    def twoSum(self, nums: List[int], target: int) -> List[int]:
        hash_table = {}
        for i, num in enumerate(nums):
            complement = target - num
            if complement in hash_table:
                return [hash_table[complement], i]
            hash_table[num] = i
        return [-1, -1]
```

在测试代码中，我们调用 `fibonacci(10)` 并打印结果。这将输出第10个斐波那契数，即55。

User: 帮我把上一段代码改成c++可以吗
Assistant: 这是一个简单的c++版本的代码:

```
#include <iostream>
#include <cstdlib>
using namespace std;

int main() {
    int n;
    cout << "Enter a number: ";
    cin >> n;
    cout << "The Fibonacci number for " << n << " is " << fibonacci(n) << endl;
    return 0;
}

int fibonacci(int n) {
    if (n <= 1)
        return n;
    else
        return fibonacci(n-1) + fibonacci(n-2);
}
```

代码解释

开源人才评价

代码纠错

实施方案:

- 调研和比较传统深度学习方法和大语言模型修复学生代码的效果
- 调研和部署开源大语言模型并在学术代码数据集上进行测评
- 利用提示工程和指令微调，提升大语言模型修复代码的准确率和生成提示信息的质量
- 将大语言模型部署至实训平台，进行效果测试

部分实验结果:

模型	Codefuse-34B	Codefuse-34B	Deepseek-33B	Deepseek-33B
学习方式	zero-shot	one-shot	zero-shot	one-shot
尝试次数	5	5	1	1
平均修复率	93%	85%	73%	71%

融合大语言模型的社区问答系统

服务层

- 动态更新
- 文档管理
- 数据库语言查询
- 对话查询
- 信息收集
- 实时展示
- 计算机领域知识检索
- PolarDB社区问答

文档管理

- 非结构化数据
- 同步维护
- 历史收集
- 向量库存储

数据统计

系统迭代

大模型问答

- 通用知识检索
- 外部知识检索
- 反馈机制

PolarDB问答

模型	模型参数	bleu4	余弦相似度	最优次数
baichuan	13B	0.033	0.728	46
ChatGLM2	6B	0.033	0.768	25
Qwen	14B	0.08	0.803	5
ChatGPT	175B	0.035	0.644	17

SQL代码生成

method	easy (163)	medium (298)	hard (114)	extra (90)	all (665)
chatglm2-6b	0.160	0.040	0.061	0.000	0.068
baichuan-13b-chat	0.129	0.067	0.044	0.011	0.071
vicuna-13b-v1.5	0.061	0.017	0.009	0.000	0.024
Qwen-14B-Chat	0.497	0.292	0.263	0.100	0.311

智能助理

实时问答

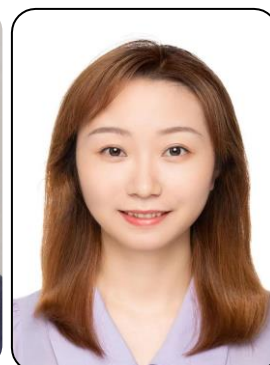
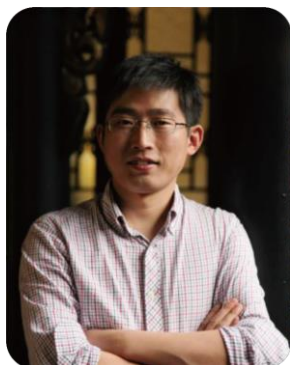
多轮问答

模型	BLEU1	BLEU2	BLEU3	BLEU4
chatglm2-6b	0.143	0.0516	0.0289	0.0198
chatglm3-6b	0.141	0.0532	0.0296	0.0197
vicuna-13b-v1.5	0.143	0.0531	0.0298	0.0196
baichuan-13b-chat	0.157	0.0644	0.0364	0.0240
Qwen-14B-Chat	0.180	0.0815	0.0489	0.0342

计算机常识问答

模型	无知识库查询时间	有知识库查询时间	余弦相似度	BLEU4
chatglm2-6b	13.8838	9.79087	0.6291	0.0104
Qwen-14B-Chat	12.3876	12.5260	0.6659	0.0180
baichuan-13b-chat	16.0946	15.2760	0.6319	0.0085
vicuna-13b-v1.5	10.6538	10.0477	0.6238	0.0084
chatglm3-6b	11.8166	12.9287	0.6225	0.0090

团队与分工



王伟
(教授)
计算教育学
开源数字生态

高明
(教授)
知识图谱
基础大模型

胡文心
(教授级高工)
大模型评测
AI4DL

陆雪松
(副教授)
代码大模型
AI4Coding

兰韵诗
(副教授)
计算语言学
AI4OSS

蒲鹏
(高级工程师)
智能运维
AI4DS

刘垚
(副教授)
高性能计算
AI4Infra

张琰彬
(工程师)
AI DevOps
AI4SE

苏斌
(助理工程师)
数字课本
AI4Infra

Open source way for GPT4ALL



OpenEduTech

README.md

{ Hello! OpenEduTech! 🌟 }

开源精神促合作、技术革新助教育、让教育不再孤单!

OpenEduTech

开源 教育

路漫漫其修远兮，We将上下而求索

View as: Public

Popular repositories

EduTechResearch 云原生、计算教育、软件过程与软件流程研究 ☆ 18 📄 9	EduPerf Benchmark suit for online education-- ☆ 7 📄 1
AccelerateDevOps Accelerate DevOps with GitHub 书籍翻译仓库 ☆ 6 📄 20	OpenDaSE 以开源的模式建立一整套的数据专业课程资源-- ☆ 5

Top languages

- Python
- Go
- HTML
- Jupyter Notebook

X-lab2017

Overview Repositories 71 Discussions Projects 2 Packages Teams 7 People 47 Settings

X-lab

开源点亮未来!

179 followers China https://github.com/X-lab2017/open... wwang@dase.ecnu.edu.cn

Unfollow

View as: Public

README.md

{ Hello! X-lab! 🌟 }

X-lab 开放实验室

{ 技术传播文明, 践行同步世界, 开源点亮未来 }

License CC BY 4.0 Xlab Discussions Xlab Newsletter Profile views 7,002

View all

Invite someone

GPT4ALL 仓库

- <https://github.com/OpenEduTech/GPT4ALL/issues/1>

开设此 GPT4ALL 仓库的原因 #1

will-ww opened this issue last week · 0 comments

will-ww commented last week

根据近期大模型赋能全民数字素养与技能提升【[GPT-DL4ALL.pdf](#)】的总结与规划，GPT4ALL 将作为一个核心的基础设施发挥作用，特创建该仓库进行讨论~

教育场景

AI4Coding AI4SE AI4DS AI4OSS

代码级生成 → 任务级生成 → 场景级生成

关键技术

测评与优化: 测评规范, 测评工具, 优化方法

边缘与轻量: AIGC一体机, 适配工具箱, BigDL项目

GPT4ALL: 提示词工程, 低代码平台

基础设施

Meta LLaMA 2, HuggingFace, Stanford Alpaca, OpenAI GPT, 川, 云, 百度, 腾讯

NVIDIA, AMD, intel, Cambium, Arista

教育场景	教	学	练	测	创	评
AI能力域						
基础知识 (AI4DL)	• 教案生成 • 出题组卷 • 知识图谱	• 自适应学习 • 智能助教	• 个性化练习 • 自动纠错 • 拓展练习	• 自动批改 • 个性化试题 • 自动打标签	• 创意生成 • 项目教学 • 智能查新	• 测评大模型 • 能力画像
编程技能 (AI4Coding)	• 编程知识图谱	• 代码助教	• 代码纠错	• 代码自动出题 • 代码自动打靶	—	• 代码能力评测
工程能力 (AI4SE)	• 工程知识图谱	• DevGPT	• AIOps 助教	• DevOps 评测	• 工程场景生成 • 工程管理 Agent	• 工程能力评测
数据分析 (AI4DS)	• 数据科学图谱	• GPT-DB • GPT-GDS	• 实训数据生成 • 分析代码生成	• 分析模型评测 • SQL/GQL生成	• 数据作品辅助 • 数据场景生成	• SQL/GQL评测
开源协作 (AI4OSS)	• 开源知识图谱 • 开源项目标签	—	• GitOps助教	• 开源社区问答	• 开源协作机器人	• 贡献度评价 • 开源人才画像
基础设施 (AI4Infra)	• 基础代码大模型、代码大模型评测、超轻量级大模型 (BigDL)、教育场景低代码平台、大模型一体机					

相关文档:

- [关于开展 22-23 级同学研究课题深入与项目实践的促进贴](#) X-lab2017/open-research#256



Ecosystem

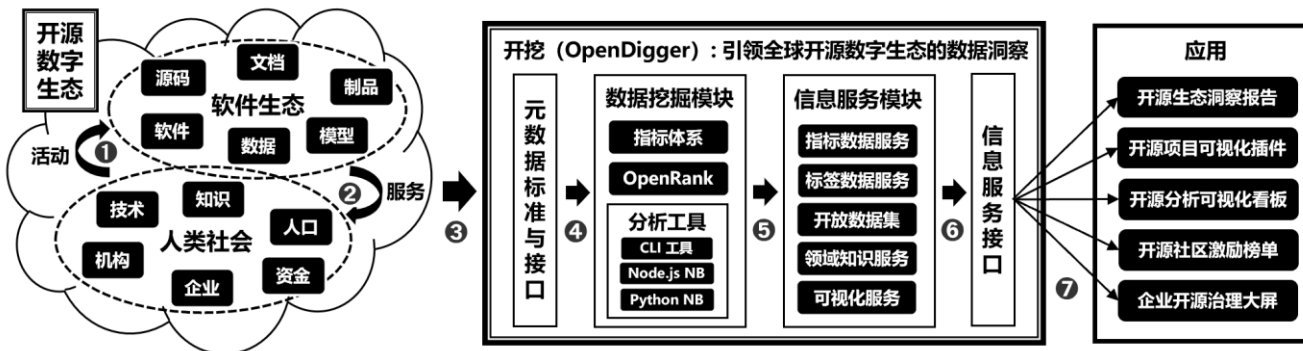
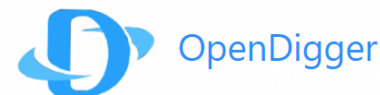
图例

- Existing
- 2023
- 2024



AOSC Leaderboard					XSOSI Leaderboard		OpenAtom Leaderboard		ChangeFundation Leaderboard	
教学		竞赛		学术		标准		内容		
OSS101		OpenSODA		OpenPerf		ONES		OpenBook		
SOS		OpenTEA		OSSCon		OpenWanderland / OpenResearch				
OpenLeaderboard			OpenDigger (OpenRank)				OpenDashboard			
Hypercrx		OSGraph		指标	模型	开放数据集		OpenShare		
OpenGalaxy			标签	工具	数据服务接口		OpenCertified			
数据基础设施与数据工程平台 (本地私有化)										
ClickHouse		MongoDB		Neo4j		TuGraph		TiDB	其他 DB	
GitHub		libraries.io		Gitee		GitLab		AtomGit	GitLink	
Stack Overflow		Slack		Twitter / 微博		Google/百度		arXiv		

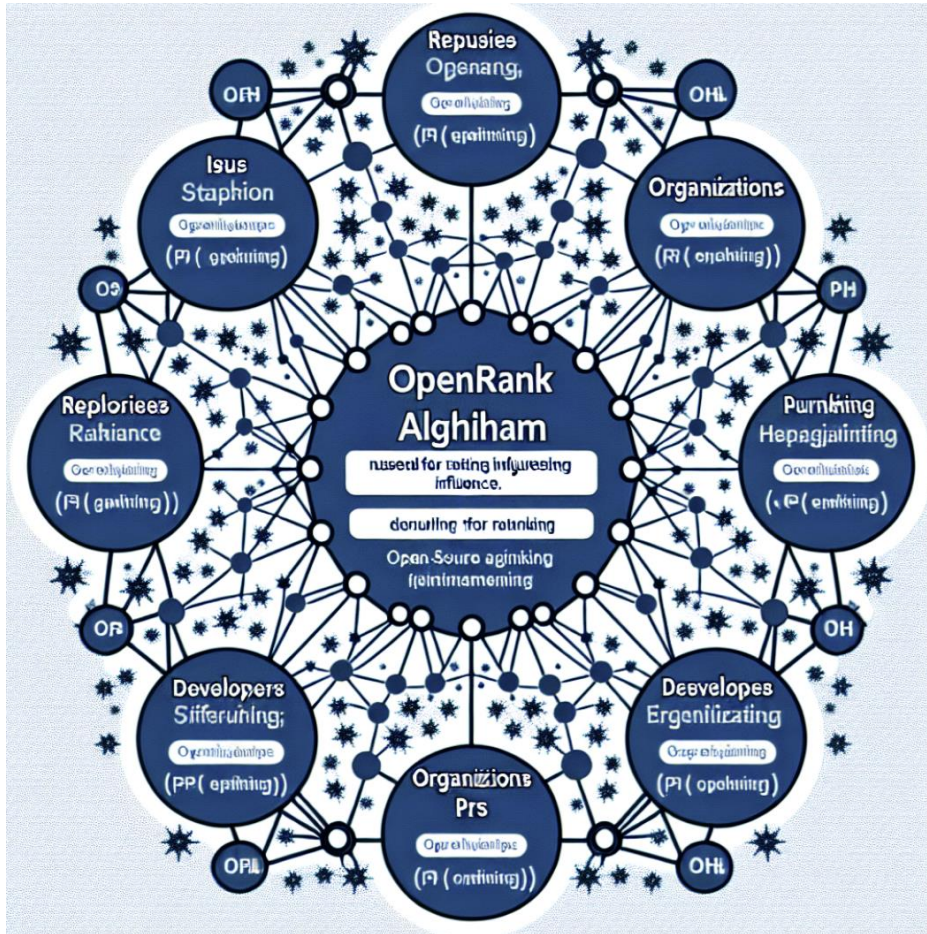
OpenDigger



OpenDigger 是一个面向开源协作数字生态的一站式数据挖掘与信息服务项目，目标是构建开源领域的数字基础设施，促进开源生态的持续发展。所支持的开源领域数据包括 GitHub 和 Gitee 行为日志数据、制品库数据、安全漏洞数据、社区问答数据等。包括 Hypercrx、OpenLeaderboard、OpenGalaxy 等均是其下游开源项目。

- GitHub 行为日志数 58 亿条
- Gitee 行为日志数 1600 万条
- NPM/PyPI 等制品库数据 620 万条
- CVE 安全漏洞数据 16 万条
- StackOverflow 问答帖 2500 万条
- 标签类数据，其中 GitHub 组织数量 413 个，涉及仓库数 89427 个

OpenRank



OpenRank Leaderboard: Motivating Open Source Collaborations Through Social Network Evaluation in Alibaba

Shengyu Zhao¹, Xiaoya Xia²
 Brian Fitzgerald³, Xiaozhou Li⁴, Valentina Lenarduzzi¹, Davide Taibi¹
 Rong Wang⁵, Wei Wang², Chunqi Tian¹

¹Tongji University, China – ²East China Normal University, China
³Lero - Irish Software Research Centre, Ireland – ⁴University of Oulu, Finland – ⁵Alibaba Group, China
 frank_zsy@tongji.edu.cn; xiaoya@stu.ecnu.edu.cn; b@lero.ie;
 xiaozhou.li@oulu.fi; valentina.lenarduzzi@oulu.fi; davide.taibi@oulu.fi;
 tunan.wr@alibaba-inc.com; wwang@dase.ecnu.edu.cn; tianchunqi@tongji.edu.cn

ABSTRACT
 Open source has revolutionized how software development is carried out, with a growing number of individuals and organizations contributing to open source projects. As the importance of open source continues to grow, companies also expect to grow thriving and sustainable open source communities with continued contributions and better collaborations. In this study, we applied the contribution leaderboard to seven open source projects initiated by Alibaba. We conducted a case study to investigate the perceptions and facts regarding how to motivate collaboration through gamification. Specifically, we employed a social network algorithm, OpenRank, to evaluate and steer developers' contributions. We validated the effectiveness of OpenRank by comparing it with other evaluation metrics and surveying developers. Through semi-structured interviews and project metric analysis, we found that the OpenRank Leaderboard can promote transparent communication environments, a better community atmosphere, and improved collaboration behavior.

KEYWORDS
 Open source contribution, Social network, Leaderboard, Gamification

ACM Reference Format:
 Shengyu Zhao¹, Xiaoya Xia² and Brian Fitzgerald³, Xiaozhou Li⁴, Valentina Lenarduzzi¹, Davide Taibi¹, Rong Wang⁵, Wei Wang², Chunqi Tian¹, 2024. OpenRank Leaderboard: Motivating Open Source Collaborations Through Social Network Evaluation in Alibaba. In *Proceedings of 46th International Conference on Software Engineering (ICSE 2024)*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/nmmmmn.nmmmm>

¹Xiaoya Xia and Shengyu Zhao contributed equally to this research.
²Wei Wang is the corresponding author.

1 INTRODUCTION
 The increasing popularity of open source projects has led to the development of an ecosystem in which individuals and organizations can collaborate to create high-quality software that is available to all [1]. Unlike traditional software development activities that are conducted internally with designated tasks for specific individuals, open source development relies on voluntary work that requires extensive asynchronous communication and distributed collaboration [2]. With the rise of companies as significant users and even major players in the open source software ecosystem, they are increasingly taking the initiative to create, open source, and maintain their own projects as a way to contribute to the community and build technical influence in the industry [3]. Naturally, companies want their open source projects to flourish as thriving communities, receiving contributions from non-employees rather than just from their own employees. This requires attracting new developers to projects in the first place (magnetism) [4, 5] and subsequently retaining these developers over time (stickiness) [6–8]. To achieve this, factors such as continued contribution, open collaboration, and fostering a positive community atmosphere are essential for long-term sustainability [9].

While some previous work has studied the role of gamification in motivating open source contributions, this has involved students [10]. In this study, we present a real-world industrial practice implemented by Alibaba to encourage contributions to its open source projects. Our approach involves constructing a network algorithm that evaluates the contributions made by open source developers. Additionally, we leverage a monthly updated leaderboard to incentivize contributions and foster a sense of healthy competition among developers.

While previous work has explored methods for evaluating the efforts of open source contributors [11–13], to our knowledge none have approached the task from a social network perspective. The collaborative nature of open source development has long



Everything You Need To Know About OpenRank

- 文档**
 - [OpenDigger 技术报告](#)
 - [如何评价一个开源项目 \(一、活跃度\)](#)
 - [如何评价一个开源项目 \(二、影响力\)](#)
 - [如何评价一个开源项目 \(三、价值流网络\)](#)
 - [开源码力榜背后的算法模型](#)
 - [开放协作的世界里，每一份贡献都值得回报](#)
 - [基于 OpenRank 的开源项目内开发者贡献评价](#)
 - [OpenRank 算法实现优化历程](#)
- 视频**
 - [OpenRank：一种开源社区贡献计算方法](#)
 - [OpenRank精讲 \(上\)](#)
 - [OpenRank精讲 \(下\)](#)
 - [OpenRank 应用：Node 开源项目排名分析](#)
 - [OpenRank 应用：NPM 开源生态网络分析](#)
- 论文**
 - [OpenRank position paper](#)
 - [OpenDigger position paper](#)
 - [OpenPerf position paper](#)
- 大赛**
 - [开放原子2023 OpenSODA 大赛](#)

OpenLeaderboard



https://open-leaderboard.x-lab.info/ 最近更新于 2023年12月12日

目标: 企业 项目 开发者 指数: 活跃度 OpenRank

地区: 中国 全球 时间: Month 2023/11 详情:

排名	企业	活跃度
1	Microsoft	59520.25 ↘ 398.85
2	Huawei	25456.9 ↗ 11375.43
3	Google	23733.75 ↘ 1381.13
4	Amazon	22783.52 ↘ 3538.49
5	Meta	9866.89 ↘ 1024.67
6	Alibaba	7760.73 ↗ 740.76
7	Elastic	6582.63 ↘ 145.92
8	HashiCorp	6395.06 ↘ 555.06
9	VMWare	5201.1 ↘ 266
10	Ant group	4920.02 ↗ 455.95

目标: 企业 项目 开发者 指数: 活跃度 OpenRank

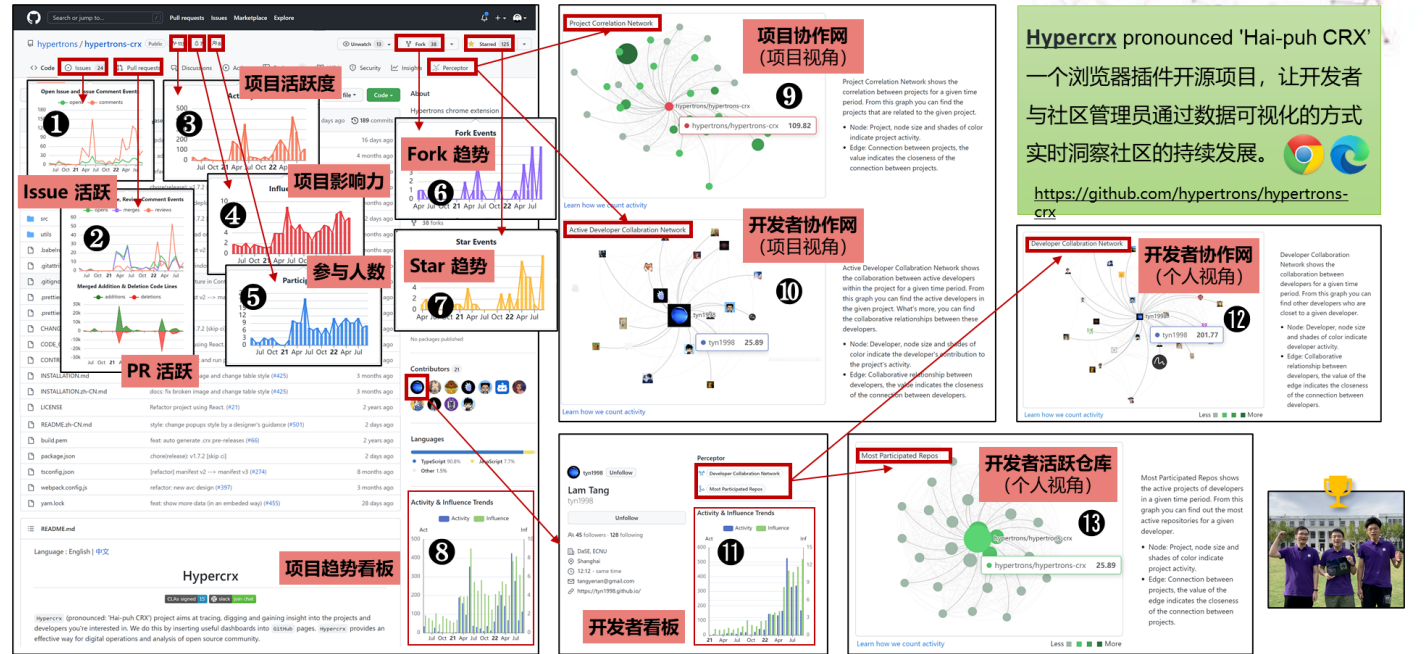
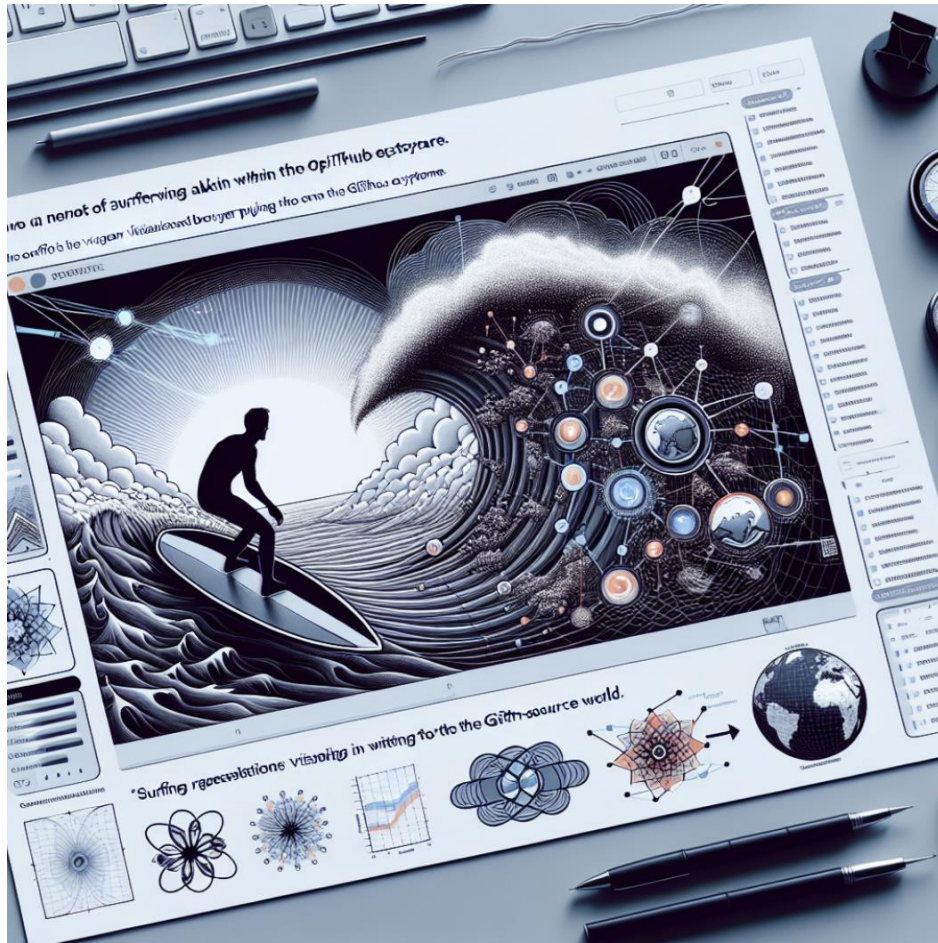
地区: 中国 全球 时间: Month 2023/11

排名	企业	OpenRank
1	Microsoft	41150.65 ↗ 390.86
2	Google	15249.99 ↘ 151.65
3	Amazon	12677.07 ↘ 157.67
4	Huawei	8292.41 ↗ 1176.10
5	Meta	7046.47 ↘ 129.02
6	Elastic	5522.04 ↗ 64.60
7	Alibaba	4648.66 ↘ 65.23
8	HashiCorp	4397.07 ↘ 78.36
9	VMWare	3771.57 ↗ 4.69
10	Intel	3602.01 ↗ 7.19

Powered by OpenDigger Mulan Community

2022年 6 月北美 Open Source Summit OSPO 上正式发布~

Hypercrx



Hypercrx pronounced 'Hai-puh CRX'
 一个浏览器插件开源项目，让开发者
 与社区管理员通过数据可视化的方式
 实时洞察社区的持续发展。
<https://github.com/hypertrons/hypertrons-crx>

Hypercrx: A browser extension for insights into GitHub projects and developers
 Anonymous Author(s)



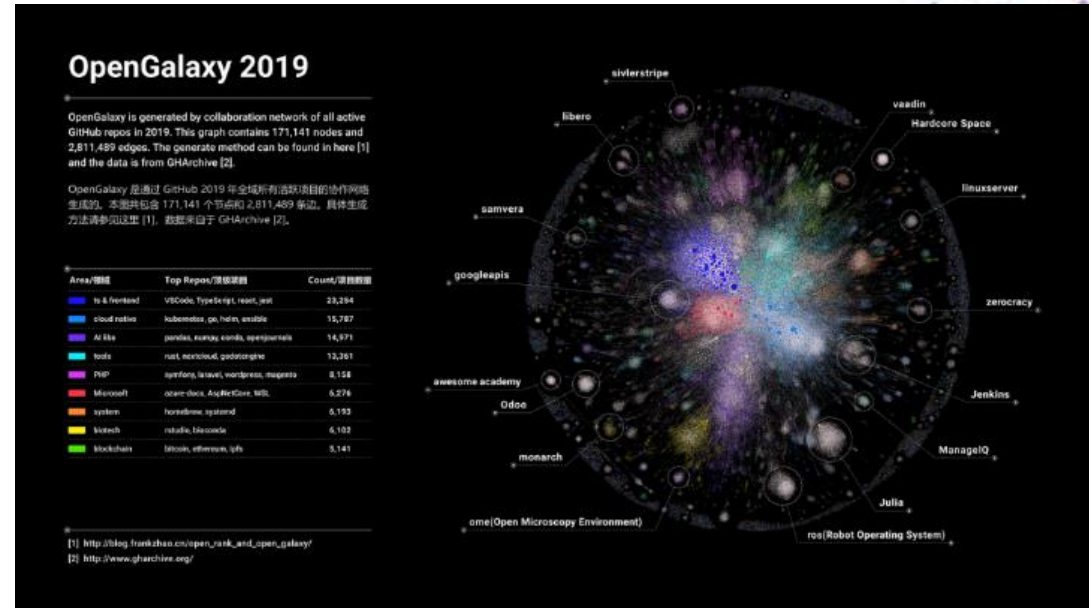
The 32nd IEEE/ACM International Conference on Program Comprehension
 Lisbon, Portugal, April 15-16, 2024

ISSPP (GPT) Seeking for GPT-powered docs support for your proje...
 [OSPP 2023] Project 1: Implement the new "Repo Collections" E...
 [OSPP 2023] Project 2: Design and implement several variable ...

Filter: 39 Open / 398 Closed
 Author: Label: Projects: Milestones: Assignees: Sort: +

- [Feature] Add contributors metric into charts. [Add/Remove](#)
- [Meeting] Hypercrx Biweekly Meeting: 2024-01-03
- [Bug] racing bar graph differs heavily at beginning of every month (2k+ data)
- [Feature] Improvement Suggestion for the xAxis of the Issues Tooltip Chart

OpenGalaxy



OpenGalaxy: An interactive exploration platform for a visualized GitHub Full Domain collaboration network

ABSTRACT

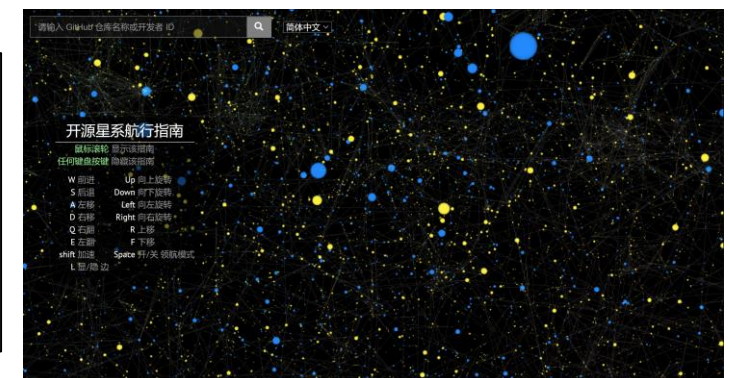
In this work, we introduce OpenGalaxy - an interactive exploration platform for a visualized GitHub Full Domain collaboration network based on 3D force-oriented layouts. We first collected GitHub domain-wide log data, built a developer-repository heterogeneous collaboration network, calculated both the influence value for each repository and the activity value for each developer, finally per-

both developers and repositories. In order to quantify the contributions of developers, we introduced an activity metric, and for measuring the project's influence, we employed a weighted PageRank algorithm. Finally, a 3D force-guided layout was implemented to effectively visualize this intricate collaborative network. Furthermore, we implemented an interactive 3D game-like mobile mode to enrich user exploration. This tool not only supports dynamic

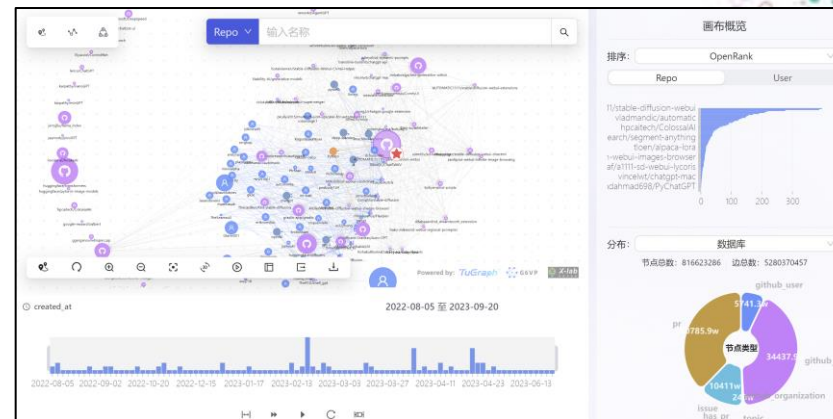
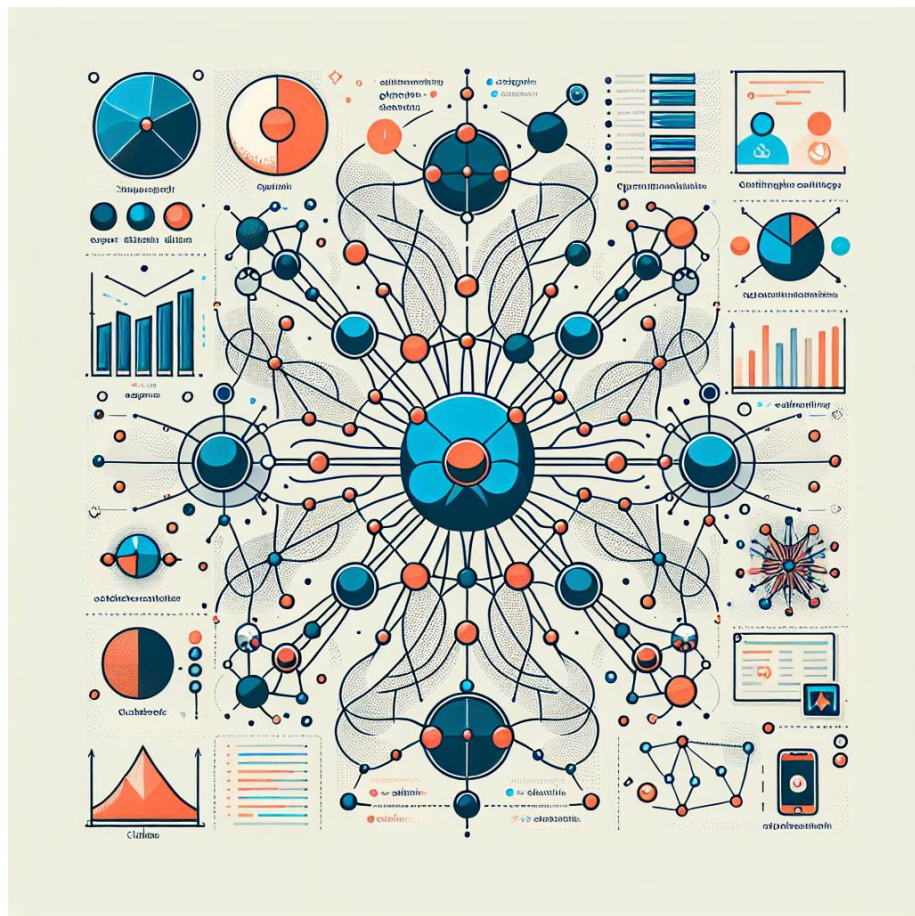


The 32nd IEEE/ACM International Conference on Program Comprehension

Lisbon, Portugal, April 15-16, 2024



OSGraph



OpenSODA



OpenDigger 开源软件生态数据分析挖掘平台挑战赛

历时6个多月，总共125支队伍参加，最终16支队伍分享了50万元的大赛奖金。

一等奖作品

name	match
x.Lab2027/open-dig	2023-08
x.Lab2027/open-dig	2023-08
x.Lab2027/open-dig	2023-09
voetis/curve	2019-10
voetis/curve	2019-11
voetis/curve	2019-12
voetis/curve	2020-01
voetis/curve	2020-02

获奖作品

(1) 任务类 (T)

- T1: **OpenDigger 指标的实现与优化** (开发类任务, 预估难度: ★)
- T2: **命令行交互的指标结果查询子模块** (开发类任务, 预估难度: ★★)
- T3: **基于机器学习的指标拟合与优化** (数据科学类任务, 预估难度: ★★★)

(2) 作品类 (W)

- W1: **可视化艺术作品或洞察报告** (数据创意类作品, 预估难度: ★)
- W2: **开源领域的可视化大屏** (Dashboard类作品, 预估难度: ★★)
- W3: **开源协作网络可视化** (网络可视化类作品, 预估难度: ★★★)

OpenPerf



第??卷 第??期 计算机学报 Vol.?? No.??
20??年??月 CHINESE JOURNAL OF COMPUTERS ??? 20??

OpenPerf: 面向开源生态可持续发展的数据科学基准测试体系

韩凡宇¹⁾, 毕枫林²⁾, 赵生宇³⁾, 李晖³⁾, 游兰⁴⁾, 王伟¹⁾, 钱卫宁¹⁾, 周傲英¹⁾

¹⁾ (华东师范大学数据科学与工程学院 上海 200062)
²⁾ (同济大学电子与信息工程学院 上海 200070)
³⁾ (贵州大学计算机科学与技术学院 贵阳 550025)
⁴⁾ (湖北大学计算机与信息工程学院 武汉 430062)

摘要 基准测试是指通过设计科学的测试方法、测试工具和测试系统, 实现对一类测试对象的某项性能指标进行定量的和可对比的测试。随着人工智能时代的到来, 诸如 ImageNet、DataPerf 等这类新型的 AI 基准测试数据逐步成为学术界和工业界的共识性标准。然而, 开源领域数据类型多样性、研究问题的广泛性以及协作网络的复杂性使得构建开源基准测试面临较大的挑战。当前, 关于开源生态的研究大多基于某一项具体的研究点展开分析而缺少对开源生态基准体系的构建, 为了解决开源领域“有数无基”的局面, 本文提出一种面向开源生态可持续发展的数据科学基准测试体系 (OpenPerf), 它定义了开源研究领域的 9 个基准测试任务, 其中涵盖了时间序列数据、文本数据和图数据 3 种数据类型; 研究问题包括回归问题、分类问题、推荐问题、排序问题、网络构建问题和异常检测问题共 6 种; 根据提出的基准测试任务, 实现了 3 项典型的任务类基准测试结果、2 项指数类基准以及 1 项标杆类基准, 其中 2 项指数类基准被中国电子技术标准化研究院作为开源社区治理的评估标准; OpenPerf 同时也作为一个开源项目, 以基准测试即服务 (Benchmarking as a Service, BaaS) 的形式给学术界、工业界、基金会等不同组织提供服务, 本文最后通过 3 个应用在阿里、蚂蚁以及华东师范大学等国内知名公司和高校的实际案例验证了 OpenPerf 在推动开源生态健康发展中所起到的关键作用。

关键词 基准测试; 开源生态; 可持续发展; 基准任务; 应用案例
中图法分类号 TP391

OpenPerf
Open Source
HAN Fan-Yu
WU Lan

¹⁾ (School of Data Science)
²⁾ (School of Electronic Information)
³⁾ (College of Computer Science and Technology)
⁴⁾ (School of Computer Science and Technology)

收稿日期: 2024-01-15; 最终修改稿收到日期: 2024-04-05; 2024-05-15 录用
基金项目: 国家自然科学基金 (62372458, 62372459, 62372460, 62372461, 62372462, 62372463, 62372464, 62372465, 62372466, 62372467, 62372468, 62372469, 62372470, 62372471, 62372472, 62372473, 62372474, 62372475, 62372476, 62372477, 62372478, 62372479, 62372480, 62372481, 62372482, 62372483, 62372484, 62372485, 62372486, 62372487, 62372488, 62372489, 62372490, 62372491, 62372492, 62372493, 62372494, 62372495, 62372496, 62372497, 62372498, 62372499, 62372500, 62372501, 62372502, 62372503, 62372504, 62372505, 62372506, 62372507, 62372508, 62372509, 62372510, 62372511, 62372512, 62372513, 62372514, 62372515, 62372516, 62372517, 62372518, 62372519, 62372520, 62372521, 62372522, 62372523, 62372524, 62372525, 62372526, 62372527, 62372528, 62372529, 62372530, 62372531, 62372532, 62372533, 62372534, 62372535, 62372536, 62372537, 62372538, 62372539, 62372540, 62372541, 62372542, 62372543, 62372544, 62372545, 62372546, 62372547, 62372548, 62372549, 62372550, 62372551, 62372552, 62372553, 62372554, 62372555, 62372556, 62372557, 62372558, 62372559, 62372560, 62372561, 62372562, 62372563, 62372564, 62372565, 62372566, 62372567, 62372568, 62372569, 62372570, 62372571, 62372572, 62372573, 62372574, 62372575, 62372576, 62372577, 62372578, 62372579, 62372580, 62372581, 62372582, 62372583, 62372584, 62372585, 62372586, 62372587, 62372588, 62372589, 62372590, 62372591, 62372592, 62372593, 62372594, 62372595, 62372596, 62372597, 62372598, 62372599, 62372600, 62372601, 62372602, 62372603, 62372604, 62372605, 62372606, 62372607, 62372608, 62372609, 62372610, 62372611, 62372612, 62372613, 62372614, 62372615, 62372616, 62372617, 62372618, 62372619, 62372620, 62372621, 62372622, 62372623, 62372624, 62372625, 62372626, 62372627, 62372628, 62372629, 62372630, 62372631, 62372632, 62372633, 62372634, 62372635, 62372636, 62372637, 62372638, 62372639, 62372640, 62372641, 62372642, 62372643, 62372644, 62372645, 62372646, 62372647, 62372648, 62372649, 62372650, 62372651, 62372652, 62372653, 62372654, 62372655, 62372656, 62372657, 62372658, 62372659, 62372660, 62372661, 62372662, 62372663, 62372664, 62372665, 62372666, 62372667, 62372668, 62372669, 62372670, 62372671, 62372672, 62372673, 62372674, 62372675, 62372676, 62372677, 62372678, 62372679, 62372680, 62372681, 62372682, 62372683, 62372684, 62372685, 62372686, 62372687, 62372688, 62372689, 62372690, 62372691, 62372692, 62372693, 62372694, 62372695, 62372696, 62372697, 62372698, 62372699, 62372700, 62372701, 62372702, 62372703, 62372704, 62372705, 62372706, 62372707, 62372708, 62372709, 62372710, 62372711, 62372712, 62372713, 62372714, 62372715, 62372716, 62372717, 62372718, 62372719, 62372720, 62372721, 62372722, 62372723, 62372724, 62372725, 62372726, 62372727, 62372728, 62372729, 62372730, 62372731, 62372732, 62372733, 62372734, 62372735, 62372736, 62372737, 62372738, 62372739, 62372740, 62372741, 62372742, 62372743, 62372744, 62372745, 62372746, 62372747, 62372748, 62372749, 62372750, 62372751, 62372752, 62372753, 62372754, 62372755, 62372756, 62372757, 62372758, 62372759, 62372760, 62372761, 62372762, 62372763, 62372764, 62372765, 62372766, 62372767, 62372768, 62372769, 62372770, 62372771, 62372772, 62372773, 62372774, 62372775, 62372776, 62372777, 62372778, 62372779, 62372780, 62372781, 62372782, 62372783, 62372784, 62372785, 62372786, 62372787, 62372788, 62372789, 62372790, 62372791, 62372792, 62372793, 62372794, 62372795, 62372796, 62372797, 62372798, 62372799, 62372800, 62372801, 62372802, 62372803, 62372804, 62372805, 62372806, 62372807, 62372808, 62372809, 62372810, 62372811, 62372812, 62372813, 62372814, 62372815, 62372816, 62372817, 62372818, 62372819, 62372820, 62372821, 62372822, 62372823, 62372824, 62372825, 62372826, 62372827, 62372828, 62372829, 62372830, 62372831, 62372832, 62372833, 62372834, 62372835, 62372836, 62372837, 62372838, 62372839, 62372840, 62372841, 62372842, 62372843, 62372844, 62372845, 62372846, 62372847, 62372848, 62372849, 62372850, 62372851, 62372852, 62372853, 62372854, 62372855, 62372856, 62372857, 62372858, 62372859, 62372860, 62372861, 62372862, 62372863, 62372864, 62372865, 62372866, 62372867, 62372868, 62372869, 62372870, 62372871, 62372872, 62372873, 62372874, 62372875, 62372876, 62372877, 62372878, 62372879, 62372880, 62372881, 62372882, 62372883, 62372884, 62372885, 62372886, 62372887, 62372888, 62372889, 62372890, 62372891, 62372892, 62372893, 62372894, 62372895, 62372896, 62372897, 62372898, 62372899, 62372900, 62372901, 62372902, 62372903, 62372904, 62372905, 62372906, 62372907, 62372908, 62372909, 62372910, 62372911, 62372912, 62372913, 62372914, 62372915, 62372916, 62372917, 62372918, 62372919, 62372920, 62372921, 62372922, 62372923, 62372924, 62372925, 62372926, 62372927, 62372928, 62372929, 62372930, 62372931, 62372932, 62372933, 62372934, 62372935, 62372936, 62372937, 62372938, 62372939, 62372940, 62372941, 62372942, 62372943, 62372944, 62372945, 62372946, 62372947, 62372948, 62372949, 62372950, 62372951, 62372952, 62372953, 62372954, 62372955, 62372956, 62372957, 62372958, 62372959, 62372960, 62372961, 62372962, 62372963, 62372964, 62372965, 62372966, 62372967, 62372968, 62372969, 62372970, 62372971, 62372972, 62372973, 62372974, 62372975, 62372976, 62372977, 62372978, 62372979, 62372980, 62372981, 62372982, 62372983, 62372984, 62372985, 62372986, 62372987, 62372988, 62372989, 62372990, 62372991, 62372992, 62372993, 62372994, 62372995, 62372996, 62372997, 62372998, 62372999, 62373000, 62373001, 62373002, 62373003, 62373004, 62373005, 62373006, 62373007, 62373008, 62373009, 62373010, 62373011, 62373012, 62373013, 62373014, 62373015, 62373016, 62373017, 62373018, 62373019, 62373020, 62373021, 62373022, 62373023, 62373024, 62373025, 62373026, 62373027, 62373028, 62373029, 62373030, 62373031, 62373032, 62373033, 62373034, 62373035, 62373036, 62373037, 62373038, 62373039, 62373040, 62373041, 62373042, 62373043, 62373044, 62373045, 62373046, 62373047, 62373048, 62373049, 62373050, 62373051, 62373052, 62373053, 62373054, 62373055, 62373056, 62373057, 62373058, 62373059, 62373060, 62373061, 62373062, 62373063, 62373064, 62373065, 62373066, 62373067, 62373068, 62373069, 62373070, 62373071, 62373072, 62373073, 62373074, 62373075, 62373076, 62373077, 62373078, 62373079, 62373080, 62373081, 62373082, 62373083, 62373084, 62373085, 62373086, 62373087, 62373088, 62373089, 62373090, 62373091, 62373092, 62373093, 62373094, 62373095, 62373096, 62373097, 62373098, 62373099, 62373100, 62373101, 62373102, 62373103, 62373104, 62373105, 62373106, 62373107, 62373108, 62373109, 62373110, 62373111, 62373112, 62373113, 62373114, 62373115, 62373116, 62373117, 62373118, 62373119, 62373120, 62373121, 62373122, 62373123, 62373124, 62373125, 62373126, 62373127, 62373128, 62373129, 62373130, 62373131, 62373132, 62373133, 62373134, 62373135, 62373136, 62373137, 62373138, 62373139, 62373140, 62373141, 62373142, 62373143, 62373144, 62373145, 62373146, 62373147, 62373148, 62373149, 62373150, 62373151, 62373152, 62373153, 62373154, 62373155, 62373156, 62373157, 62373158, 62373159, 62373160, 62373161, 62373162, 62373163, 62373164, 62373165, 62373166, 62373167, 62373168, 62373169, 62373170, 62373171, 62373172, 62373173, 62373174, 62373175, 62373176, 62373177, 62373178, 62373179, 62373180, 62373181, 62373182, 62373183, 62373184, 62373185, 62373186, 62373187, 62373188, 62373189, 62373190, 62373191, 62373192, 62373193, 62373194, 62373195, 62373196, 62373197, 62373198, 62373199, 62373200, 62373201, 62373202, 62373203, 62373204, 62373205, 62373206, 62373207, 62373208, 62373209, 62373210, 62373211, 62373212, 62373213, 62373214, 62373215, 62373216, 62373217, 62373218, 62373219, 62373220, 62373221, 62373222, 62373223, 62373224, 62373225, 62373226, 62373227, 62373228, 62373229, 62373230, 62373231, 62373232, 62373233, 62373234, 62373235, 62373236, 62373237, 62373238, 62373239, 62373240, 62373241, 62373242, 62373243, 62373244, 62373245, 62373246, 62373247, 62373248, 62373249, 62373250, 62373251, 62373252, 62373253, 62373254, 62373255, 62373256, 62373257, 62373258, 62373259, 62373260, 62373261, 62373262, 62373263, 62373264, 62373265, 62373266, 62373267, 62373268, 62373269, 62373270, 62373271, 62373272, 62373273, 62373274, 62373275, 62373276, 62373277, 62373278, 62373279, 62373280, 62373281, 62373282, 62373283, 62373284, 62373285, 62373286, 62373287, 62373288, 62373289, 62373290, 62373291, 62373292, 62373293, 62373294, 62373295, 62373296, 62373297, 62373298, 62373299, 62373300, 62373301, 62373302, 62373303, 62373304, 62373305, 62373306, 62373307, 62373308, 62373309, 62373310, 62373311, 62373312, 62373313, 62373314, 62373315, 62373316, 62373317, 62373318, 62373319, 62373320, 62373321, 62373322, 62373323, 62373324, 62373325, 62373326, 62373327, 62373328, 62373329, 62373330, 62373331, 62373332, 62373333, 62373334, 62373335, 62373336, 62373337, 62373338, 62373339, 62373340, 62373341, 62373342, 62373343, 62373344, 62373345, 62373346, 62373347, 62373348, 62373349, 62373350, 62373351, 62373352, 62373353, 62373354, 62373355, 62373356, 62373357, 62373358, 62373359, 62373360, 62373361, 62373362, 62373363, 62373364, 62373365, 62373366, 62373367, 62373368, 62373369, 62373370, 62373371, 62373372, 62373373, 62373374, 62373375, 62373376, 62373377, 62373378, 62373379, 62373380, 62373381, 62373382, 62373383, 62373384, 62373385, 62373386, 62373387, 62373388, 62373389, 62373390, 62373391, 62373392, 62373393, 62373394, 62373395, 62373396, 62373397, 62373398, 62373399, 62373400, 62373401, 62373402, 62373403, 62373404, 62373405, 62373406, 62373407, 62373408, 62373409, 62373410, 62373411, 62373412, 62373413, 62373414, 62373415, 62373416, 62373417, 62373418, 62373419, 62373420, 62373421, 62373422, 62373423, 62373424, 62373425, 62373426, 62373427, 62373428, 62373429, 62373430, 62373431, 62373432, 62373433, 62373434, 62373435, 62373436, 62373437, 62373438, 62373439, 62373440, 62373441, 62373442, 62373443, 62373444, 62373445, 62373446, 62373447, 62373448, 62373449, 62373450, 62373451, 62373452, 62373453, 62373454, 62373455, 62373456, 62373457, 62373458, 62373459, 62373460, 62373461, 62373462, 62373463, 62373464, 62373465, 62373466, 62373467, 62373468, 62373469, 62373470, 62373471, 62373472, 62373473, 62373474, 62373475, 62373476, 62373477, 62373478, 62373479, 62373480, 62373481, 62373482, 62373483, 62373484, 62373485, 62373486, 62373487, 62373488, 62373489, 62373490, 62373491, 62373492, 62373493, 62373494, 62373495, 62373496, 62373497, 62373498, 62373499, 62373500, 62373501, 62373502, 62373503, 62373504, 62373505, 62373506, 62373507, 62373508, 62373509, 62373510, 62373511, 62373512, 62373513, 62373514, 62373515, 62373516, 62373517, 62373518, 62373519, 62373520, 62373521, 62373522, 62373523, 62373524, 62373525, 62373526, 62373527, 62373528, 62373529, 62373530, 62373531, 62373532, 62373533, 62373534, 62373535, 62373536, 62373537, 62373538, 62373539, 62373540, 62373541, 62373542, 62373543, 62373544, 62373545, 62373546, 62373547, 62373548, 62373549, 62373550, 62373551, 62373552, 62373553, 62373554, 62373555, 62373556, 62373557, 62373558, 62373559, 62373560, 62373561, 62373562, 62373563, 62373564, 62373565, 62373566, 62373567, 62373568, 62373569, 62373570, 62373571, 62373572, 62373573, 62373574, 62373575, 62373576, 62373577, 62373578, 62373579, 62373580, 62373581, 62373582, 62373583, 62373584, 62373585, 62373586, 62373587, 62373588, 62373589, 62373590, 62373591, 62373592, 62373593, 62373594, 6

GPT4ALL for DL4ALL

