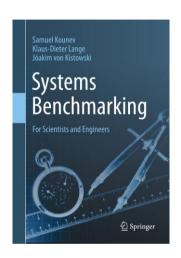
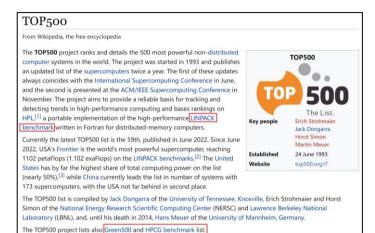


EduPerf 的愿景与目标

为什么要做 Benchmark?

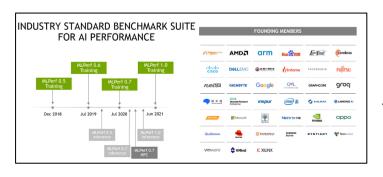






Jack Dongarra





MLPerf Al Benchmarks

为什么要做 Benchmark?



BenchCouncil: International Open Benchmark Council

Home Benchmarks Ranking Conferences Journal Organization Awards Publications Testbed

Challenges ComputerCouncil Users CN

Top Level Projects

- BigDataBench
- AlBench
- ScenarioBench
- Clinical AlBench
- SAIBench

Ranking

- HPC AI500 Ranking
- Datacenter Al Ranking
- AloT Ranking

Incubator Projects

- BENCHCPU
- WPC
- PowerSystemBench
- OLxPBench

Conferences

- Bench 2022
- Bench 2021

FICC 2021

International Open Benchmark Council (BenchCouncil) is a non-profit organization that aims to benchmark and standardize emerging technologies like Big Data. Al, Chips, HPC, Metaverse, etc.

Since its founding, BenchCouncil bears four fundamental responsibilities: establish unified benchmark science and engineering across multi-disciplines (https://www.sciencedirect.com/science/article/pii/S2772485922000515); keep the benchmarks and standards community open, inclusive, and growing; define the problems or challenges of emerging and future computing; promote benchmark-based quantitative approaches to tackle multidisciplinary and interdisciplinary challenges.

What's New:

10/13/2022: TBench Special Issue of "Open-source Computer Systems": Call for Papers (Submission Site). This special issue focuses on studies in exploring the software and hardware co-design space in high-end computer systems, studies in advancing open-source movement including novel abstraction and methodology, open-source hardware, open-source software, measurement and optimization tools.

10/13/2022: Open-source Computer System (OpenCS) Workshop Call for Participation (Nov 10-11, 2022, 8:00 am UTC-5, Preliminary Program, Free registration).

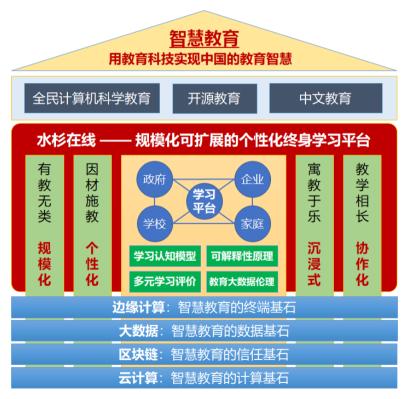
10/13/2022: Bench 2022 Call for Participation (Nov 6-11, 2022, 8:00 am UTC-5, Preliminary Program, Free registration). Highlights: BenchCouncil Achievement Award Lecture; BenchCouncil Rising Star Award Lectures; Paper presentations of Bench 2022 papers and TBench papers; Two workshops on OpenBench and Open-source Computer System (OpenCS).





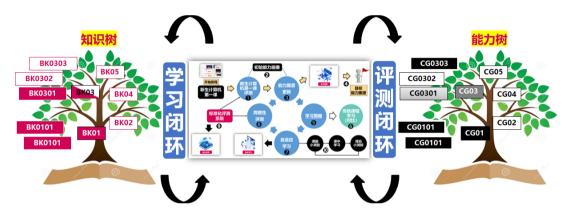


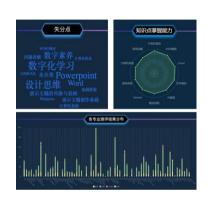
和我们有些什么关系?

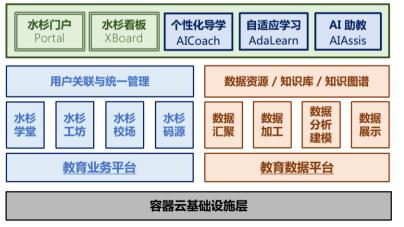


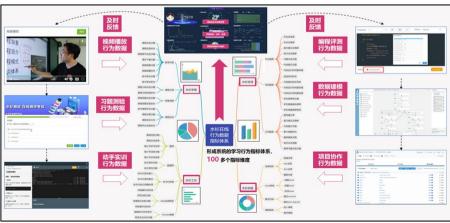


和我们有些什么关系?

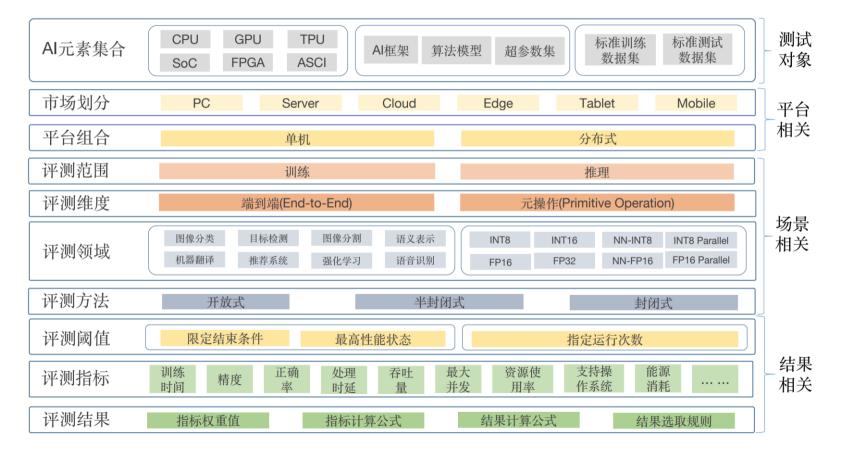




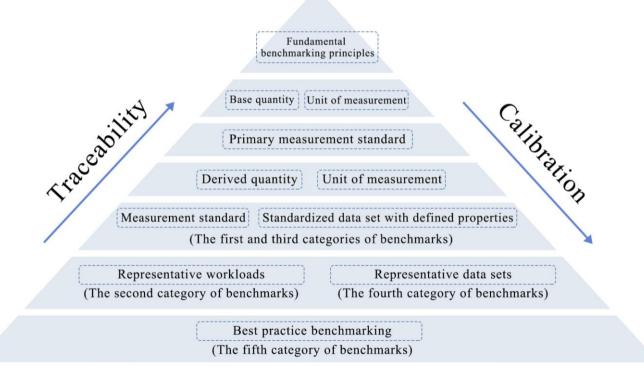




基于基准测试的研究工作分类

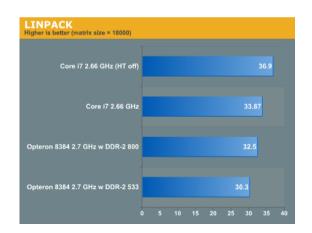


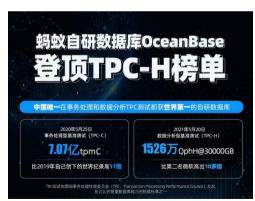
基准架构: Benchmark hierarchy (五大类)

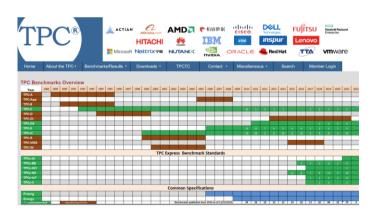


第一类:测量标准

The first category of the benchmark is a measurement standard.
 In the computer discipline, the Linpack benchmark is of this category, which is widely used to report the performance of a high-performance computer.



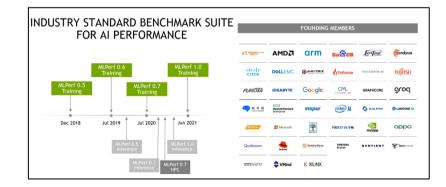


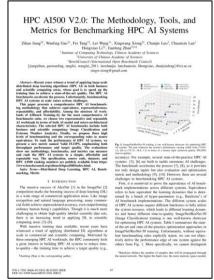


第二类: 代表性系统 workloads

The second one is the representative workloads that run on the systems under measurement. The application benchmarks or synthetic benchmarks in the computer discipline are of this

category.





第三类:数据与算法基准

 The third is a standardized data set that represents real-world data science problem, with defined properties, some of which have ground truth. The benchmark of this category is often used to measure against different algorithms. The state-of-the-art algorithm implementation plus the data set usually constitutes the benchmark of the second category.







We especially (but not exclusively) call for submissions which will contribute to at least one of the following:

Call for Papers - GLB 2022

- Real-World Datasets: Novel real-world graph-structured datasets—especially large-scale, application-oriented, and publicly accessible datasets.
- Synthetic Datasets (New): Synthetic graph-structured datasets that are well-supported by graph theory, network science, or empirical studies, and can be used to reveal limitations of existing graph learning methods.
- Software Packages: Software packages which enable streamlined benchmarking large-scale online graphs, crawling or crowdsourcing of graph data, and generation of realistic synthetic graphs.
- Tasks: New learning tasks and applications on different types of graphs, at different levels (e.g., node, edge, and (sub)graph), with a special focus on real-world and industry-oriented problems.
- . Metrics: New evaluation procedures and metrics of graph learning associated with the various tasks and datasets.
- Benchmarks: Works benchmarking multiple existing GNNs on non-trivial tasks and datasets. We explicitly encourage works
 that reveal limitations of existing models or optimize matches between network designs and problems.
- Task Taxonomy (New): Discussions towards a more comprehensive and fine-grained taxonomy of graph learning tasks.

第四类: 行业指数

The fourth is a representative data set, used as a reference. For example, a financial benchmark is an index (statistical measure), calculated from a representative set of underlying data, is used as a reference for financial instruments or contracts. Well-known financial benchmarks include the London Interbank Offered Rate (Libor) and the Euro Interbank Offered Rate

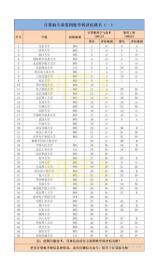




第五类: 业界标杆

The fifth is the industry best practices in different domains.
 Benchmarking is the continuous process of searching the industry best practices that lead to superior performance and measuring products, services, and processes against them.

Types of Benchmarking Competitive or Performance Benchmarking Internal Benchmarking Process Benchmarking





和我们很有关系



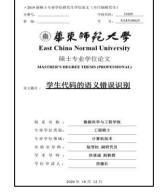




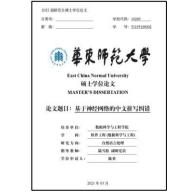


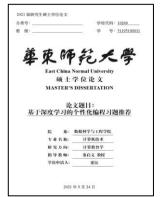












和我们很有关系



BenchCouncil: International Open Benchmark Council

Home	Benchmarks	Ranking (Conferences	Journal	Organization	Award	ds Publicatio	ns	Testbed	
Challenges	BigDataBench	uncil Users	Bench				Achievement		AI	
	AlBench		Federated				Awards		Big Data	
op Level Proje	ScenarioBench	itional Open Ber	conferences	nchCoun	cil) is a non-profit or	ganizati	Rising Star Awards	ıcł	Al & Big Data	ize
BigDataBenc	Clinical AlBench	ing technologies	Conference lists	ips, HPC	, Metaverse, etc.	- M. (1887) (1) (1) (1)	Distinguished		Cloud &	
AlBench	SAIBench	s founding, Benc	Proceedings	ındamer	tal responsibilities: es	tablish u	Doctoral	ier	Datacenters	acros
ScenarioBen Clinical AIBe	BENCHCPU	lisciplines (https://			ce/article/pii/S277248		Dissertation Award		CPU	40.00
SAIBench		ınity open, inclus	i Review rules	ne the p	roblems or challenges	of eme	Best Paper Awards	ıρι	WPC	ımark
<u>anking</u>	WPC	quantitative approaches to tackle multidisciplinary and interdisciplinary challenges and interdisciplinary challenges are supported by the control of the co							Federated learning	
HPC AI500 R	PowerSystemBench	New:					Reproducible Research		Storage	
Datacenter / AloT Ranking	OLxPBench				Computer Systems":				Energy Efficiency	

Incubator Projects

- BENCHCPU
- WPC
- PowerSystemBench
- OLxPBench

Conferences

- Bench 2022
- Bench 2021

10/13/2022: TBench Special Issue of "Open-source Computer Systems": Call for Papers (Submission Special Issue of "Open-source Computer Systems": Call for Papers (Submission Special Issue of Topen-source software and hardware co-design space in high-end computer systems, studies in advancing open-source movement including novel abstraction and methodology, open-source hardware, open-source software, measurement and optimization tools.

10/13/2022: Open-source Computer System (OpenCS) Workshop Call for Participation (Nov 10-11, 2022, 8:00 am UTC-5, Preliminary Program, Free registration).

10/13/2022: Bench 2022 Call for Participation (Nov 6-11, 2022, 8:00 am UTC-5, Preliminary Program, Free registration). Highlights: BenchCouncil Achievement Award Lecture; BenchCouncil Rising Star Award Lectures; Paper presentations of Bench 2022 papers and

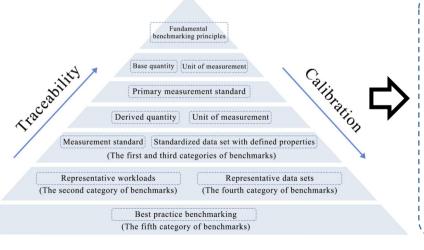
EduPerf 的几个目标

- 1. 构建一个面向水杉在线与计算教育学场景下的"教育科技基准体系"
- 2. 该体系可以包括所有五类基准,从第二类(系统)、第三类(数据/算法)做起
- 3. 将"水杉在线系统"与"计算教育学"中的场景进行抽象与定义,包括:
 - Real-World Datasets、Synthetic Datasets、Software Packages、Tasks、Metrics、 Benchmarks、Task Taxonomy
- 4. 沉淀整个教育方向实验室的工程与科研成果
- 5. 以开源的方式进行运作,从 ECNU 做起,逐步组织更多的人
- 6. 逐步建立包括国际会议、期刊、行业标准、开源项目、课程等完整的体系
- 7. 成为团队包括科研项目、工程项目、奖项申报、基地申报、人才申报等公共服务

EduPerf hierarchy (tentative)









GitHub 上的仓库

https://github.com/OpenEduTech/EduPerf

