

数据治理与数据挖掘分营简介

数据治理与数据挖掘分营将对人工智能时代的**数据治理与挖掘**展开探究。当前人工智能的蓬勃发展离不开数据。大模型的骨架是神经网络，但智能源于海量数据。提升数据质量、挖掘数据规律，对进一步提升产业智能化程度具有重要意义。分营由**浙江大学数据库与大数据分析（DAILY）实验室**，**浙江大学软件学院大数据研究中心**相关团队主办。根据 CSRankings.org 计算机学科排名最新数据，本团队在数据库领域的研究位列**世界第四位**。

CSRankings: Computer Science Rankings

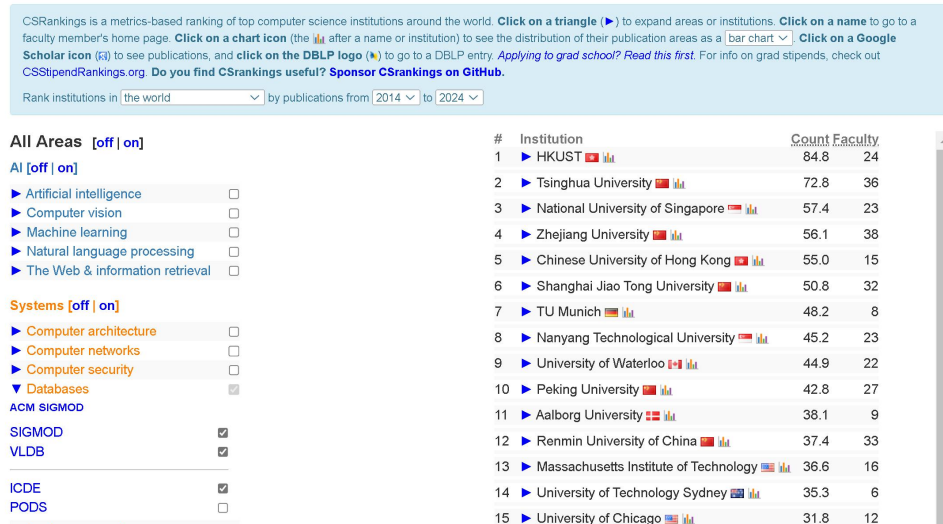


图 1 CSRankings.org 数据库领域学科排名

团队的带头人是高云君教授。高云君教授是浙江大学求是特聘教授，博士生导师，国家杰出青年科学基金获得者（2020）、国家优秀青年科学基金获得者（2015），现为 ACM SIGSPATIAL 中国分会副主席，全省大数据智能计算重点实验室主任，浙江大学软件学院常务副院长，浙江大学计算机软件研究所副所长。研究方向为数据库、大数据管理与分析、DB 与 AI 融合。已在国内外顶级学术期刊或会议 ACM Computing Surveys、TODS、VLDBJ、TKDE、TPDS、TOIS、TMC、SIGMOD、VLDB、ICDE、SIGKDD、SIGIR 等发表论文 170 余篇，出版中英文学术专著 4 部、教材 2 本（含 1 本在线教材），授权发明专利 30 余项，登记软件著作权 4 项，并获 VLDB 2024 最佳研究论文提名、DASFAA 2024 最佳论文、ICDE 2022 最佳论文、ICDE 2019 优秀论文（One of the Best Papers）、APWeb-WAIM 2018 最佳论文、SIGMOD 2015 最佳论文提名、ICDE 2015 优秀论文等。主持国家杰出青年科学基金、优秀青年科学基金、国家重点研发计划、973 计划、国家自然科学基金联合基金重点等 20 余项项目/课题，并获 2019 年度中国电子学会科技进步特等奖、2016 年度教育部科技进步一等奖、2011 年度浙江省科学技术一等奖。培养多名博/硕士生获得或入选 KDD Cup 2022 风电预测（Wind Power Forecast）赛道全球冠军、中国计算机学会（CCF）优秀博士学位论文奖、ACM SIGMOD 中国优博奖、ACM SIGSPATIAL 中国优博奖、ACM 中国优博奖提名、

浙江省优秀博士学位论文提名、中国电子学会硕士学位论文激励计划、浙江省优秀硕士学位论文、国家特聘青年专家、国家优秀青年科学基金（海外）、ACM SIGMOD 中国新星奖、ACM SIGSPATIAL 中国新星奖、浙江省杰出青年科学基金、甬江人才计划、华为天才少年等。

团队目前与阿里巴巴通义千问实验室，微软亚洲研究院，字节跳动，蚂蚁智能引擎、华为智能创新 Lab 等头部企业的研究部门建立了广泛合作关系，提供丰富的实习机会。实验室与海外高校如 UCSD, Purdue U, UNSW, NTU 等合作紧密，可提供海外交流机会。培养的软件学院硕士生，斩获 KDD Cup 2022 世界冠军，获得国家奖学 5 人次，平均发表数据库及数据挖掘领域 CCF-A 类论文 1 篇以上（包括 SIGMOD, VLDB, KDD 等），硕士同学全部拿到阿里，字节，腾讯，百度等头部企业 offer，包括蚂蚁星等 Special Offer。

本分营主要导师团队及研究方向如下，具体信息可参阅各位导师个人主页：

团队	姓名	职称	研究方向	个人主页
数据库与大数据分析实验室	高云君	教授/博导	数据库与人工智能融合	https://person.zju.edu.cn/gaoyj_cn
	毛玉仁	百人计划研究员/博导	基于大模型智能体的数据管理与分析	https://person.zju.edu.cn/maoyuren
	柯翔宇	百人计划研究员/博导	资源受限环境下的轻量化向量数据管理与应用	https://person.zju.edu.cn/xiangyuke
	房子荃	百人计划研究员/博导	时空大数据管理与智能分析	https://person.zju.edu.cn/fangziquan
	朱铁凡	百人计划研究员/博导	结合新型硬件的向量数据管理与分析	https://person.zju.edu.cn/NB24090
大数据研究中心	张微	教授/博导	基于卫星导航和时空智能的智慧城市平台开发与应用	https://person.zju.edu.cn/cstzhangwei
	贝毅君	副研究员/博导	多模态数据异常检测与分析	https://person.zju.edu.cn/beiyj
	程学林	副研究员/硕导	基于多智能体协同的数据处理与协作平台	https://person.zju.edu.cn/cxlin
	张启飞	副研究员/硕导	唇语识别、多模态等	https://person.zju.edu.cn/zhangqf

各导师研究方向详细介绍如下：

研究方向一：数据库与人工智能融合（负责老师：高云君）

在“数据驱动”和“智能引导”优势互补的指导思想下，充分发挥数据库技术在大数据存储与处理能力方面的优势，结合人工智能方法强大的建模与学习能力，促进更高层次智能的产生，建立“大数据智能”的新模式。

研究方向二：基于大模型智能体的数据管理与分析（负责老师：毛玉仁）

面向包含文本、图片、表格、序列等数据模态的多源异构数据，利用大模型智能体的规划、工具调用和决策能力，开发服务海量多源异构数据的管理和分析的专属智能体系统。

研究方向三：资源受限环境下的轻量化向量数据管理与应用（负责老师：柯翔宇）

面向移动端、边缘计算、物联网终端等资源受限环境（如受限的计算能力、内存、存储空间、网络带宽、功耗），研究高维向量数据的高效、轻量化管理与智能分析理论与技术。重点突破在严格资源约束条件下，向量数据的紧凑表征、低开销存储、自适应索引、近似查询与分布式协同分析等核心挑战，为增强现实、移动端大模型应用、智能感知终端等场景提供理论与技术支撑。

研究方向四：时空大数据管理与智能分析（负责老师：房子荃）

面向智慧城市、智能交通等场景中的时空大数据（如 GPS 轨迹、遥感影像、交通流量、传感器时序数据等），研究基于分布式平台（如 Spark、Flink）的时空数据高效存储与并行计算技术；探索多模态时空数据的表征学习与智能建模方法，构建融合领域知识的时空大模型分析系统。

研究方向五：结合新型硬件的向量数据管理与分析（负责老师：朱轶凡）

面向大规模高维向量数据，结合分布式环境、NVMe 固态硬盘、GPU 等新型硬件的强大存储与计算能力，深入探索向量数据的表征、存储、索引、查询与分析等，从而支撑大模型、云数据库等场景下的高效数据管理与智能分析。

研究方向六：基于卫星导航和时空智能的智慧城市平台开发与应用（负责老师：张微）

面向智慧城市建设中的多源异构数据（如地理空间信息、物联网传感数据、城市运行管理数据等），利用卫星导航定位、遥感地理信息、智能传感等技术，结合时空智能分析能力，构建开放化、数字化、智能化的城市综合管理平台，实现城市数据的精准采集、智能分析、科学管理和协同决策，赋能智慧城市的多场景应用。

研究方向七：多模态数据异常检测与分析（负责老师：贝毅君）

在多模态异常数据的自动发现与分析领域，当前研究聚焦于时序数据的动态特征捕捉、图结构数据的关联模式挖掘，以及音频 / 图像 / 视频等非结构化数据的语义特征提取。通过跨模态数据融合技术，实现多源异构数据的关联分析与异常模式识别，为智能监测、风险预警等场景提供更全面的决策支持。

研究方向八：基于多智能体协同的数据处理与协作平台（负责老师：程学林）

利用多个大模型智能体之间的分工与交互，结合其工具调用和智能决策能力，面向跨境物流领域，实现自动化的邮件监听、多源异构数据解析、比对、存储以及通知任务，通过多智能体分工与交互实现高效自动化，提升数据处理的准确性和协作效率。

研究方向九：唇语识别、多模态、基于国产 NPU 平台的 AI SoftPLC 系统（负责老师：张启飞）

面向工业控制、智能制造、自动驾驶等领域，开发基于国产 NPU 平台的 SoftPLC 系统及 AI 控制算法。