

浙江大学数据科学研究中心
Center for Data Science, Zhejiang University

1. 中心简介

“浙江大学数据科学研究中心”成立于 2017 年 5 月 18 日，是以统计学、应用数学、计算机科学和管理学为核心支撑学科，以大数据理论、应用研究和人才培养为主的校设学术创新研究机构。研究中心强调与经济学、医学、生命科学、社会学、工学等众多学科领域的交叉融合，在获取基础研究的硕果同时，注重于科研技术成果的转化。面向市场和产业需求进行人才培养，着重培养研究型和技术型人才，同时通过组建跨学科、跨领域的教学、研究师资队伍，大力培养复合型数据科学人才。中心发展的方向包括研究数据科学的理论与方法；为处理海量的、高维的数据发展分析的工具和算法，进而为应用和管理提供方向；设立数据科学的教学平台，培养相关领域内的国际化高端人才。研究中心致力于面向大数据战略前瞻问题和社会需求，为不同学科之间实现有渠道的、有组织合作提供平台，凝聚和培养数据科学一流人才，打造高质量、高效率的数据科学核心技术与理论研究创新平台。

研究中心基本定位：汇聚和培养数据科学高端人才，覆盖各重要学科，开展大数据原创性研究与开发应用，聚焦解决重大问题，创造具有实际意义的创新算法，切实解决中国大数据领域的前瞻性问题。

研究中心的核心目标：通过机制创新，汇聚和培养国内外数据科学领域的一流人才，构建数据科学高精尖协同创新基地。利用杭州大数据行业的优势和浙江大学多学科交叉的特点，尽快使浙江大学成为国际大数据研究领域的引领者之一。

研究中心的主要功能：

(1) 汇聚国内外高层次统计学、计算机科学、应用数学、管理科学等数据科学相关人才，建立高水平数据科学专业化研究队伍，在大数据分析，机器学习，人工智能等几个重点方向开展具有前瞻性的国际领先水平的研究，将研究中心建设为数据科学核心理论与技术的创新研究中心；

(2) 在浙江大学内部，促进跨院系协同的学科力量整合，同时加强与政府部门、高科技公司的交流合作，搭建数据科学高能级交叉研究平台；

(3) 设立数据科学教学平台，在短期内完善数据科学博士项目、建立专业硕士项目，并在长期规划发展联合硕士项目，探索会聚式研究生协同培养模式，设立本科项目，将研究中心打造为多学科融合的交叉复合创新人才培养基地。

中心主页：<https://cds.zju.edu.cn>

2. 导师团队介绍

苗晓晔，浙江大学数据科学研究中心长聘副教授、博士生导师。长期从事数据库与数据智能方向研究，重点关注不完整数据管理、数据质量治理、数据价值评估与定价、联邦数据协作等问题。

苗晓晔团队（ZJU-Data VIP Lab）聚焦 Value, Integrity, and Pricing of Data，致力于构建让数据可治理、可评估、可流通、可交易的理论方法与系统工具。团队围绕“数据要素价值释放”这一核心目标，系统开展缺失数据智能修复、任务导向数据治理、因果忠实修复、数

据价值评估、无套利数据定价以及大模型辅助数据治理等研究。团队研究强调理论机制、系统方法与真实场景应用的结合，面向数据要素流通、遥感数据资产化、多组学数据治理、联邦数据协作和大模型时代的数据质量评估等场景，探索数据从“可用”到“可信”、从“治理”到“估值”、从“资源”到“资产”的关键技术路径。团队在国内外顶级/重要学术期刊/会议 The VLDB Journal, IEEE Trans. Knowl. Data Eng., IEEE Trans. Mobile Comput., 计算机研究与发展, Briefings in Bioinformatics, VLDB, SIGMOD, KDD, AAAI, SIGIR, ICDE 等发表论文 70 余篇，授权/受理国家发明专利 20 余项（美国专利 3 项），合作出版英文学术专著（美国 Morgan & Claypool 出版社）和中文教材《数据科学前沿技术导论》（高等教育出版社，“新一代人工智能”系列教材），并有中文学术专著《数据智能定价：理论框架、核心机制与场景应用》待出版。

苗晓晔曾获人工智能学会吴文俊人工智能科技进步一等奖，ACM SIGMOD 中国新星奖（2023），受浙江省杰青（2020）项目资助。担任 CCF 高级会员，CCF 数据库专委会执行委员，CCF 服务计算专委会执行委员。曾/正担任 Nature Communications, Patterns, VLDBJ, TKDE, TKDD, TSC, Briefings in Bioinformatics 等学术期刊评审专家；KDD 2026 Area Chair; VLDB 2025, KDD 2024-26, WWW 2024-25, ICDE 2023-27, AAAI 2022-27 等会议程序委员会委员。承担浙江省自然科学基金杰出青年项目（结题优秀）、国家自然科学基金面上项目、国家自然科学基金青年项目、国家重点研发计划子课题、国家重点研发计划青年科学家项目任务、浙江省“领雁”研发攻关计划项目课题、重大校企合作横向项目等多项。

个人主页: https://person.zju.edu.cn/miaoxy_cn

3. 招生信息

目前在研课题项目主要包括：

1. **大模型辅助的数据治理与质量评估**：研究 LLM 辅助的数据清洗、补全、错误检测、质量评估与任务导向修复方法。
2. **不完整数据管理与因果忠实修复**：研究缺失、错误和不确定数据的理论建模、可恢复性分析、因果一致修复与可信推断。
3. **数据价值评估、定价与数据要素流通**：研究面向数据交易、数据资产化和数据服务的数据估值、无套利定价、收益分配与可信流通机制。
4. **联邦与多源数据协作治理**：研究隐私保护、多源协同、多模态/多组学/遥感等场景下的联邦学习、可信数据协作与智能应用。

基本要求:

1. 数据科学、计算机、人工智能、数学、统计、软件工程等相关背景；
2. 有一定的数学和编程基础；
3. 对展开以上相关课题研究感兴趣，动手能力强，英语读写熟练，责任心强。
4. 有相关科研经验者优先。