

首届全国高校数据驱动创新研究大赛

培训计划

(1) 第一次培训

时间：2017年11月30日 下午3:00~4:30

现场培训地点：北京大学图书馆304教室

网络直播地址：<http://162.105.138.115/index.php?m=live&c=index&a=lists>



表1 第一次培训内容

主持人	主要内容	培训老师
刘雅琼 (北京大学图书馆)	大赛基本情况介绍 (30分钟)：介绍大赛的基本情况，包括大赛要求、赛制赛程、注册和成果提交流程、北京大学开放数据平台等。	罗鹏程 馆员 (北京大学图书馆) 北京大学图书馆信息化与数据中心馆员，负责北京大学开放研究数据平台的建设工作，曾参与国家自然科学基金委基础研究知识库、北京大学科研管理系统等平台的建设。参与负责本次大赛的相关组织工作。
	数据挖掘方法介绍 (30分钟)：简要介绍数据挖掘的基本流程和方法。	王继民 教授 (北京大学信息管理系) 教授，博士生导师，北京大学信息管理系副主任。研究领域包括：搜索引擎、Web数据挖掘、科学评价学、信息可视化等。近几年主持国家社科基金、国家“核高基”重大科技专项子课题、以及国家发改委、教育部、北京市科委等科研课题30余项。发表学术论文50余篇；出版专著或合著《搜索引擎原理技术与系统》、《Web用户查询日志挖掘与应用》、《中国人文社科类一级学科数据分析报告》、《“一带一路”沿线国家五通指数报告》、《国民海洋意识发展指数研究报告(2016)》等6部。获得发明专利2项；获得省部级科研奖励2项。
	现场答疑 (30分钟)	

(2) 第二次培训

时间：2017年12月01日 下午 3:30~5:00

现场培训地点：北京大学图书馆 304 教室

网络直播地址：<http://162.105.138.115/index.php?m=live&c=index&a=lists>



表 2 第二次培训内容

主持人	主要内容	培训老师
赵飞 (北京大学图书馆)	中国家庭追踪调查及分析方法 (30分钟): 对中国家庭追踪调查数据(CFPS)进行介绍, 并简要介绍相关的分析方法。	吴琼 副研究员 (北京大学社会科学调查中心) 美国宾州州立大学教育与心理测量学博士、统计学硕士。现任北京大学中国社会科学调查中心副研究员, “中国家庭追踪调查”(CFPS)项目办公室主任, 主要负责 CFPS 数据管理、数据服务、与问卷设计和执行相关的数据支持工作。加入调查中心之前, 她就职于哈佛大学人口与发展研究中心, 作为该中心的量化分析师, 她的主要职能之一是分析大型调查数据。主要研究领域包括测量学方法、认知功能的影响因素、少儿发展等, 已发表 SSCI、SCI 论文 20 余篇。
	中国健康与养老追踪调查及分析方法 (30分钟): 对中国健康与养老追踪调查数据 (CHARLS) 进行介绍, 并简要介绍相关的分析方法。	陈欣欣 副研究员 (北京大学社会科学调查中心) 浙江大学管理学博士, 现任北京大学中国社会科学调查中心副研究员, 中国健康与养老追踪调查 (CHARLS) 项目主管, 曾在斯坦福大学师从 Scott Rozelle 教授从事博士后研究。2008 年以来参与了 CHARLS 的实地执行工作, 并组织实施了 中国中老年人生命历程调查、CHARLS 第三轮追踪调查和共和国初期基层经济史调查。研究兴趣集中在微观发展经济学和老年经济学。
	国家信息中心大数据发展部数据介绍 (30分钟): 介绍国家信息中心大数据发展部的开放数据。	廖尚围 项目经理 (国信宏数公司) 国信宏数公司数据采集项目经理。曾任蓬天公司 CTO, 负责陕西省、江西省地税征管系统技术架构, 具有丰富的 J2EE 项目开发实施经验。目前主要负责国信宏数公司数据采集工作, 通过设计分布式采集平台, 实施互联网结构化、非结构化数据的采集、清洗、存储。