

附录3 培训与讲座

直播地址（所有培训和讲座都通过以下地址直播，培训结束后在官网提供录播）：

<http://162.105.138.115/index.php?m=live&c=index&a=lists>

（1） 第一次：大赛培训

时间：2018年11月19日 上午9:00~11:00

现场培训地点：北京大学图书馆 304 教室



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。
	数据预处理方法 (30分钟)：介绍数据预处理的一般技术与方法，主要包括数据清理、数据集成、数据变换、数据归约、数据离散化等。	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。
	北京市政务数据资源网开放数据介绍 (30分钟)：对北京市政务数据统一开放平台——北京市政务数据资源网所开放的数据进行介绍	高文飞 高级工程师 (北京市信息资源管理中心) 北京市信息资源管理中心数据开放部项目主管，北京市大数据工作推进小组办公室成员，长期致力于政务数据资源管理及应用相关研究和实践工作，目前主要负责政务数据开放、数据汇聚共享、大数据应用等工作。
	现场答疑 (30分钟)	

(2) 第二次：数据讲座

时间：2018年11月19日 下午 14:00~15:00

现场培训地点：北京大学图书馆 304 教室

主要内容	讲座老师
人文社科数据管理与服务的研究与实践 （60分钟）：介绍国内外人文社会科学数据管理与服务，以及所参与政府和高校科学数据中心建设的一些数据平台，如上海人口数据实验室、复旦大学人文社会科学数据研究所、上海市慧源共建共享平台、国家卫计委流动人口数据平台等等，并展示一些项目的数据可视化。	殷沈琴 副研究馆员（复旦大学人文数据研究所） 复旦大学人文社会科学数据研究所科学数据中心主任、硕士生导师，上海人口数据实验室副主任、“开放数林”政府开放数据专家委员会委员。主要从事科学数据管理、政府开放数据、社会管理与社会政策等领域的研究。承担和参与十余个国家级和省部级的项目，并多次负责数据平台的规划和部署实施工作，有丰富的数据研究和落地实践经验。

(3) 第三次：数据讲座

时间：2018年11月19日 晚上 19:00~20:00

现场培训地点：北京大学图书馆 304 教室

主要内容	讲座老师
科研常用开放数据资源的查找与获取 （60分钟）：介绍网络上开放获取的各类数据资源，并结合具体案例介绍查找各类数据资源的技巧。	朱玲 馆员（北京大学图书馆） 北京大学图书馆信息化与数据中心数据管理服务主管，负责北京大学图书馆多个系统平台项目：资源发现系统“未名学术搜索”、电子资源导航系统、北京大学开放研究数据平台、电子资源使用监控与统计系统等。已发表图情类CSSCI 论文 9 篇。

(4) 第四次：大赛培训

时间：2018年11月21日 上午10:00~11:20

现场培训地点：北京大学图书馆 304 教室

主持人	主要内容	培训老师
张元俊 (北京大学图书馆)	中国家庭追踪调查介绍 (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介绍 (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，现任北京大学中国健康与养老追踪调查 (CHARLS) 执行主任，相继组织实施了中国中老年人生命历程调查、共和国初期基层经济史调查、京津居民家庭结构和生命历程调查、CHARLS 第三轮和第四轮常规追踪调查。研究兴趣集中在微观发展经济学和老年经济学。
	专题介绍 (20 分钟)：对专题一和专题二进行详细解读和介绍。	吴琼 副研究员 (北京大学社会科学调查中心) 赵银霞 软件工程师 (北京大学社会科学调查中心) 毕业于郑州大学计算机科学与技术专业，目前从事 PHP 软件开发，WEB 前端开发。

(5) 第五次：大赛培训+数据讲座

时间：2018年11月23日 上午8:30~10:00

现场培训地点：阿卜杜勒·阿齐兹国王公共图书馆北京大学分馆地下一层报告厅

主要内容	培训（讲座）老师
<p>【培训】中国创新创业数据库介绍（30分钟）：对中国企业创新创业数据库进行介绍，该数据库集合了微观企业在创业、投资，以及企业创新产出等方面的信息。</p>	<p>杨奇明 董事长（古德科技有限公司）</p> <p>杨奇明拥有浙江大学管理学博士学位，北京大学应用经济学博士后。主持博士后基金项目和国家社科青年项目，在《经济研究》、《管理世界》等学术期刊发表论文10余篇。自北京大学从事博士后研究期间（2013.7-2016.1）开始，一直从事企业大数据的融合与开发工作，协助合作导师组织构建了朗润-龙信创新创业指数，作为联合主编出版了《中国区域创新创业报告》。</p>
<p>【讲座】Making your data count! - Why you need to care about sharing data and how we can help you（60分钟）：Following the recent Research Data Survey to find out more about Data Management and attitudes, we will introduce the importance of data for research purposes, why efficient data management is important for research to have an impact and how we can provide help.</p>	<p>Arend Küster / 柯安德（Springer Nature 开放科研大中华区总监）</p> <p>Arend 负责 Springer Nature 旗下所有开放获取业务在中国的发展，包括 Nature Communications、BMC 系列期刊、Springer Open 系列期刊等，他在科技和医学出版领域拥有非常丰富的经验，致力于促进中国学者与国际出版业科研理念的交流。</p> <p>Grace Baynes（Springer Nature 研究数据与新产品研发副总裁）</p> <p>Grace 在 Springer Nature 领导着开放数据与优秀数据实践事业。她还负责广受赞誉、高速发展的 Nature 旗下期刊 Scientific Data 的出版。她还是国际开放科研政策与实践的专家。最近，Grace 领导着开放数据团队，与 Wellcome Trust 和德国马普学会签署了开拓性的合作。</p>