

数据素养教育的设计与实施

张久珍

北京大学信息管理系

2019-8-21

鄂尔多斯



主要内容

1. 《信息素养概论》里强化数据素养有关内容
2. 参与一流高校数据服务支撑体系建设
3. 开设《数字人文》课程，聚焦数据能力



一、强化数据素养有关内容

- 图书馆学研究和实践的对象，从传统的印本文献，到电子文献。然后呢？
- 数据！数据是未来。

二、参与一流高校数据服务支撑体系建设

□ 我们看到了很多图书馆的开放数据平台，正在服务于教学与科研。



北京大学 开放研究数据平台
Peking University Open Research Data



关于

支持

联系

注册

登录

Switch to English

北京大学开放研究数据平台

重用数据，再现科研。
提升引用，跟踪影响。

开放数据
为您的研究点赞！

创建我的数据空间

或

浏览平台数据



北京大学 开放研究数据平台
Peking University Open Research Data



关于

支持

联系

注册

登录

Switch to Eng



北京大学中国调查数据资料库



中国家庭追踪调查 (CFPS)



中国健康与养老追踪调查
(CHARLS)



北京大学中国国情研究中心



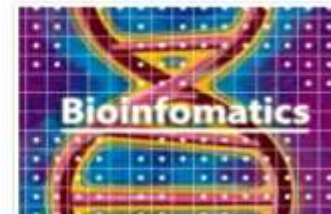
北京大学健康老龄与发展研究
中心



综合型语言知识库 (CLKB)



北京大学可视化与可视分析研究
组



学中心



信息官理系



	A	B	C	D	E
1	数据集名称	地址			
2					
3	中国历代人物传记资料库（2017）	http://opendata.pku.edu.cn/dataset/			
4					
5	工商企业登记注册信息	http://opendata.pku.edu.cn/dataset/			
6	地图兴趣点POI（Point of Interest）数据	http://opendata.pku.edu.cn/dataset/			
7	政府招投标项目数据	http://opendata.pku.edu.cn/dataset/			
8	互联网招聘网站数据	http://opendata.pku.edu.cn/dataset/			
9	互联网新闻、论坛文本数据	http://opendata.pku.edu.cn/dataset/			
10	电商商品及网店数据	http://opendata.pku.edu.cn/dataset/			
11	上市公司财报数据	http://opendata.pku.edu.cn/dataset/			
12	发明专利数据	http://opendata.pku.edu.cn/dataset/			
13	北京市政务数据资源网数据	http://opendata.pku.edu.cn/dataset/			
14	佰职公开职位大数据信息	http://opendata.pku.edu.cn/dataset/			
15	清博大数据数据源—19个国家的微博数据	http://opendata.pku.edu.cn/dataset/			
16	清博大数据数据源—19个国家的微信数据	http://opendata.pku.edu.cn/dataset/			



-
- ☐ 有没有？
 - ☐ 有多少？
 - ☐ 有没有人用？
 - ☐ 有没有人教？
 - ☐ 会有多少人会用？
 - ☐ 有多大影响力？

Deposit and share your data. Get academic credit.

Harvard Dataverse is a repository for research data. Deposit data and code here:

90,175 datasets 8,057,841 downloads

Add a dataset +

Organize datasets and gather metrics in your own repository.

A dataverse is a container for all your datasets, files, and metadata.

3,330 dataverses

Add a dataverse +

Find data across research fields, preview metadata, and download files

Search over 90,100 datasets...

Q Find

Browse by subject

Agricultural Sciences 1,262

Arts and Humanities 799

Astronomy and Astrophysics 513

Business and Management 428

Chemistry 183

Computer and Information Science 919

Earth and Environmental Sciences 1,465

Engineering 403

Law 274

Mathematical Sciences 199

Medicine, Health and Life Sciences 2,995

Physics 514

Social Sciences 38,632

Share, archive, and get credit for your data. Find and cite data across all research fields.

Search this dataverse...

 Find

[Advanced Search](#)

☒  **Dataverses (3,329)**

☒  **Datasets (90,176)**

☐  **Files (528,624)**

Dataverse Category

Research Project (1,116)

Researcher (1,019)

Organization or Institution (301)

Research Group (166)

Journal (83)

[More...](#)

Metadata Source

Harvested (60,537)

Harvard Dataverse (32,968)

Publication Year

2015 (15,446)

2011 (9,548)

2012 (8,147)

2018 (5,087)

2016 (4,073)

[More...](#)

1 to 10 of 93,505 Results

[Replication Data for: Inequities in consistent condom use among sexually experienced undergraduates in mainland China: implications for planning interventions](#)



2019-7-30

Wang, Junfang, 2019, "Replication Data for: Inequities in consistent condom use among sexually experienced undergraduates in mainland China: implications for planning interventions", <https://doi.org/10.7910/DVN/GBFCTK>, Harvard Dataverse, V1, UNF:6:Q27TrDOmtQbZLIDVn1kV3Q== [fileUNF]

Table 1 Variables in Model of consistent condom use Variables in Model Variables in data set of this study Name Definition Depend...
Consistent condom use (Y) Consistent Inconsistent=0, Consistent=1 Predisposing factors Gender (X1) Gender Female=0, Male=1 Grad...

[Replication Data for: Do Voters Polarize When Radical Parties Enter Parliament?](#)

2019-7-30 - [American Journal of Political Science \(AJPS\) Dataverse](#)



Bischof, Daniel; Wagner, Markus, 2019, "Replication Data for: Do Voters Polarize When Radical Parties Enter Parliament?", <https://doi.org/10.7910/DVN/DZ1NFG>, Harvard Dataverse, V3, UNF:6:5P0XysqluWq7I+OLFcCa/w== [fileUNF]

Do voters polarize ideologically when radical views gain political legitimacy, or does the rise of radical voices merely reflect societal...
argue that the elite polarization as signalled by radical parties' first entrance into parliament leads to voter divergence. Imme...

[Outlook 2060 U.S. by County Population Annual Estimates and Projections, 2010-2060](#)

2019-7-30 - [Harvard Library E-Resources Licensed Data Dataverse](#)



Glimpse, Warren, 2019, "Outlook 2060 U.S. by County Population Annual Estimates and Projections, 2010-2060", <https://doi.org/10.7910/DVN/FTKO40>, Harvard Dataverse, V1



Replication data for: The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Field Experiment



2014-8-24 - Review of Economics and Statistics Dataverse

Cai, Hongbin; Chen, Yuyu; Fang, Hanming; Zhou, Li-An, 2014, "Replication data for: The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Field Experiment", <https://doi.org/10.7910/DVN/27174>, Harvard Dataverse, V1, UNF:5:iNhyV0oaQqWAylfd31fceg== [fileUNF]

Cai, Hongbin, Chen, Yuyu, Fang, Hanming, and Zhou, Li-An, (2015) "The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Field Experiment." Review of Economics and Statistics 97:2, 287-300.

Author Affiliation: **Peking** University

Replication data for: One Child Policy and the Rise of Man-Made Twins



2015-8-19 - Review of Economics and Statistics Dataverse

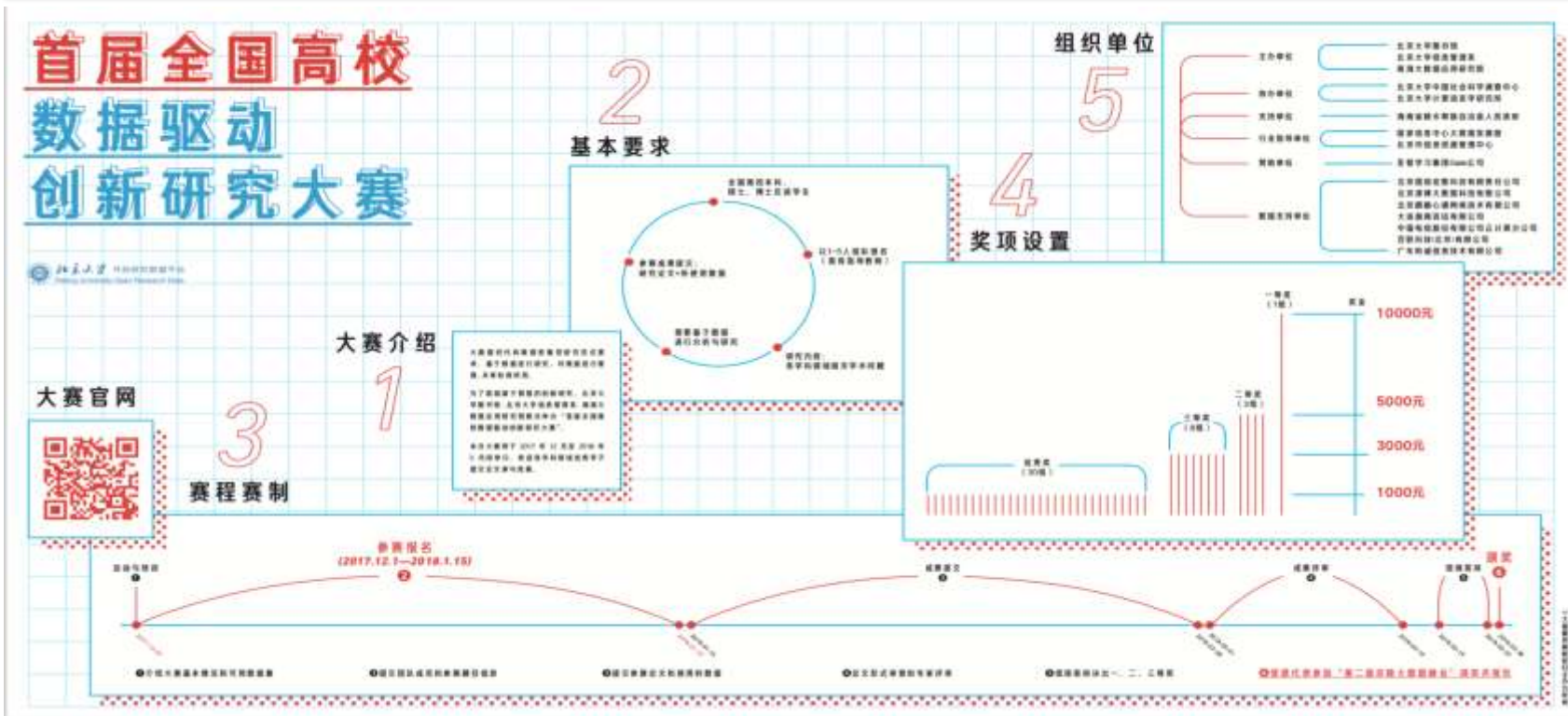
Huang, Wei; Lei, Xiaoyan; Zhao, Yaohui, 2015, "Replication data for: One Child Policy and the Rise of Man-Made Twins", <https://doi.org/10.7910/DVN/UIDQYS>, Harvard Dataverse, V1, UNF:6:6CIXkHEkWm4blAaFEq/x7w== [fileUNF]

Huang, Wei, Lei, Xiaoyan, and Zhao, Yaohui, (2016) "One Child Policy and the Rise of Man-Made Twins." Review of Economics and Statistics 467-476.

Author Affiliation: **Peking** University

2017年11月24日-2018年4月3日

赛制介绍



大赛组织与赛制

参赛形式与要求

- (1) 全国高校本科、硕士、博士在读学生均可参赛;
- (2) 以1~5人组队报名(每人只能参与一支队伍);
- (3) 要求有指导教师;
- (4) 研究内容为各学科领域相关学术问题;
- (5) 需要基于数据进行研究, 包含针对数据的相关分析和结论;
- (6) 参赛成果提交的形式为研究论文, 同时提供所使用研究数据;
- (7) 入围决赛的参赛团队, 要求参加现场答辩;
- (8) 论文和数据需要满足如下论文和数据要求;
- (9) 参赛者允许组织方对参赛作品汇集成册、展示和宣传。

论文要求

论文模板

数据要求

本
硕
博



大赛组织与赛制

大赛章程(详细)



特等奖	1	奖金20000元
一等奖	1	奖金10000元
二等奖	3	奖金5000元
三等奖	8	奖金3000元
优秀奖	120	奖金1000元

大赛简介

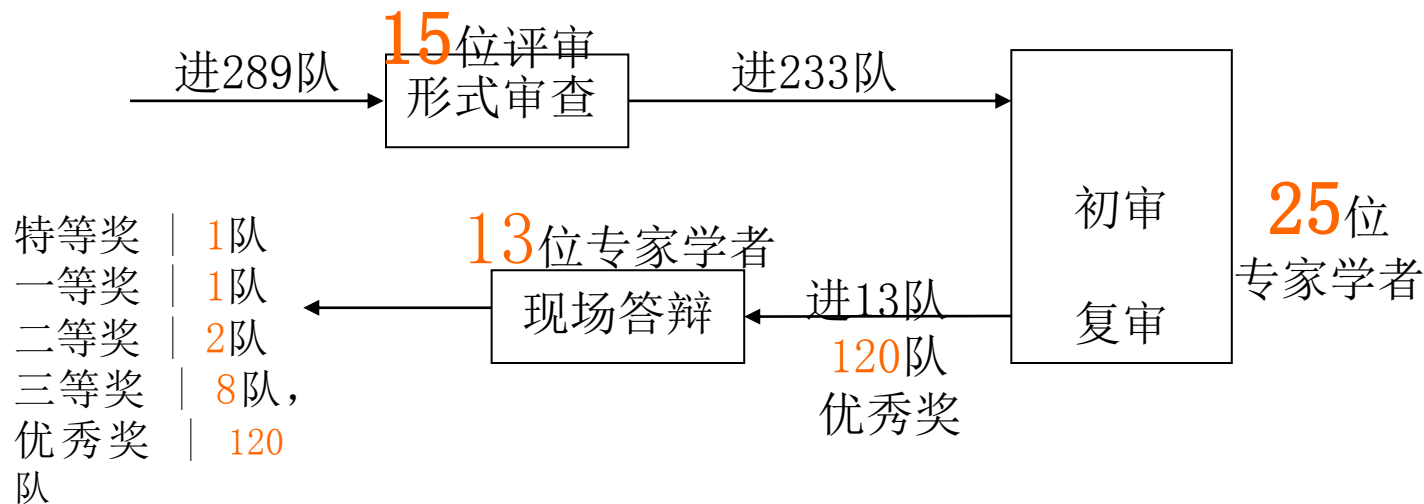
随着大数据时代和数据密集型研究范式的到来,基于数据进行研究,对数据进行管理、共享和再利用,成为学术研究的新趋势。为了鼓励各学科领域学子基于数据进行创新研究,促进研究数据的保存和共享,由国家信息中心大数据发展部、北京市信息资源管理中心作为行业指导单位,北京大学图书馆、北京大学信息管理系、南海大数据应用研究院,联合北京大学中国社会科学调查中心、北京大学计算语言学研究所、重庆市仙桃大数据与物联网职业培训学院,面向全国高校在读学生,开展首届全国高校数据驱动创新研究大赛。

本次大赛将于2017年12月至2018年3月期间举行,欢迎各学科领域优秀学子提交论文参与竞赛。

2018年4月3日 现场答辩

赛制介绍

成果评审



评判标准

论文选题

论文创新性

论文论证

论文工作量

论文规范性

原创数据

大赛成果



- 加快平台建设：国际一流数据平台，汇集更多数据
- 汇集产业大数据，服务高校课程、研究项目
- 推动高校数据驱动研究范式
- 北京大学、发改委、企业等多方合作，促进教学、科研与创新研究孵化

H135

Fe 14387

H	I	J	K	L	M	N	O	P	Q	R	S
10844	231	崔1523021	13820	本科生组	郝1河北工业大经济学	应用经济学	家庭消费行为视角下的金融资产配置	应用经济学	家庭消费行为视角下的金融资产配置	本文基于北京大学中国社	
19606	240	张120721	18810	本科生组	黄1中国人民大经济学	应用经济学	基于机器学习的“一带一路”沿线国家投资区位智能化推荐	应用经济学	基于机器学习的“一带一路”沿线国家投资区位智能化推荐	本文采取机器学习的方法	
12108	266	杨156001	15600	本科生组	何1中国人民大经济学	应用经济学	网络舆情下的人民币汇率变动——基于新浪微博大数据的研究	应用经济学	网络舆情下的人民币汇率变动——基于新浪微博大数据的研究	本文基于新浪微博大数据	
15770	279	赵1zhaozh1	18811	本科生组	侯1北京林业大经济学	应用经济学	影响中老年人住院报销情况因素 —— 基于 CHARLS数据的研究	应用经济学	影响中老年人住院报销情况因素 —— 基于 CHARLS数据的研究	【目的】了解我国中老年	
51079	17	程1chengz1	18908	研究生组	程1西南交通大管理学	公共管理	交通基础设施有助于减贫吗？——来自CFPS数据的证据	公共管理	交通基础设施有助于减贫吗？——来自CFPS数据的证据	基于中国家庭追踪调查	
15184	75	史1redwin1	15221	研究生组	殷1复旦大学 管理学	管理科学	基于机器学习的我国城市“双创”水平的评价研究	管理科学	基于机器学习的我国城市“双创”水平的评价研究	本文依托科技创新相关	
16403	103	李1354261	15140	研究生组	翟1大连海事大管理学	管理科学	北京市开放政府数据是权威的数据源吗？——对www.bjdata.gov	管理科学	北京市开放政府数据是权威的数据源吗？——对www.bjdata.gov	【目的/意义】我国各级政	
18778	196	张1187421	18742	本科生组	高1东北财经大管理学	管理科学	引入新闻短文本的个股走势预测模型	管理科学	引入新闻短文本的个股走势预测模型	大量研究已证明新	
27663	199	李1984961	15927	研究生组	李1武汉大学政管理学	公共管理	基于社会舆情与机器学习的群租房精准治理研究	公共管理	基于社会舆情与机器学习的群租房精准治理研究	【目的/意义】作为城市	
17940	246	梅1seekie1	13641	本科生组	张1东华大学 管理学	管理科学	我国农产品电商生态系统分析	管理科学	我国农产品电商生态系统分析	农产品电子商务作为农	
22439	252	徐1xuhany1	15542	研究生组	田1东北财经大管理学	管理科学	相对剥夺、社会资本与老年心理健康	管理科学	相对剥夺、社会资本与老年心理健康	在我国老龄化日益严重	
15577	264	明1ei_min1	18811	本科生组	韩1北京林业大管理学	管理科学	基于知识图谱的人才需求分析——以互联网行业为例	管理科学	基于知识图谱的人才需求分析——以互联网行业为例	【目的/意义】随着互联网	
13540	286	高1584771	13006	本科生组	吴1武汉大学信管理学	管理科学	招聘职位与求职者匹配优化研究——基于“58同城”	管理科学	招聘职位与求职者匹配优化研究——基于“58同城”	【目的/意义】互联网时代	
10666	302	张1595911	17854	本科生组	杜1济南大学教管理学	管理科学	基于求职网站大数据的可视化研究分析	管理科学	基于求职网站大数据的可视化研究分析	数据挖掘和数据分析的	
11686	362	徐1247451	13060	本科生组	赵1四川大学 管理学	工商管理	地区产业政策的影响 ——以“一带一路”倡议背景下西安产业发展	工商管理	地区产业政策的影响 ——以“一带一路”倡议背景下西安产业发展	本文通过对西安市2012-	
11592	390	方1201531	15167	本科生组	李1武汉大学信管理学	管理科学	社交媒体情境下基于机器学习技术的 药物不良反应研究	管理科学	社交媒体情境下基于机器学习技术的 药物不良反应研究	生物医药技术和制药技	
21228	450	张1877951	17816	本科生组	张1浙江大学公管理学	公共管理	问世间情为何物？——基于CHARLS数据研究丧偶对于抑郁的	公共管理	问世间情为何物？——基于CHARLS数据研究丧偶对于抑郁的	随着我国人口老龄化和	
14943	458	周1280921	18707	本科生组	赵1武汉大学信管理学	管理科学	中美数据科学人才技能需求及培养方案探究	管理科学	中美数据科学人才技能需求及培养方案探究	【目的/意义】随着信息科	
12755	465	王1131631	15516	本科生组	高1武汉大学经管理学	管理科学	众筹项目融资决策建议——基于融资效率影响因素分析	管理科学	众筹项目融资决策建议——基于融资效率影响因素分析	众筹作为快速兴起	
11028	489	吴1150001	18801	本科生组	严1北京大学政管理学	公共管理	中国网民的社会公共事务参与行为与态度研究——基于线上线下	公共管理	中国网民的社会公共事务参与行为与态度研究——基于线上线下	本研究的主旨是探究中	
13012	291	孙1sdq101	18693	本科生组	陈1兰州大学经经济学	应用经济学	集聚经济理论下企业选址的抉择	应用经济学	集聚经济理论下企业选址的抉择	本文的目的在于探讨在	
15403	313	严1yanyay1	17682	研究生组	段1南开大学金经济学	应用经济学	捕捉“隐形的篮子”——微观经济智能决策系统	应用经济学	捕捉“隐形的篮子”——微观经济智能决策系统	金融投资强调分散风险	
13853	322	魏1193991	15871	本科生组	丁1湖北工业大经济学	应用经济学	从招聘需求中挖掘学历、经验、公司性质对薪酬的影响-以金融行	应用经济学	从招聘需求中挖掘学历、经验、公司性质对薪酬的影响-以金融行	利用2017年招聘职位招	
15144	334	李1leefutu1	18025	本科生组	蔡1暨南大学经经济学	应用经济学	经济发展下多元教育需求的涌现 ——来自中国华德福教育的证据	应用经济学	经济发展下多元教育需求的涌现 ——来自中国华德福教育的证据	改革开放以来，中	

大赛成果

121所高校 | **968** 人 | **289** 组队伍成功提交了参赛作品。

- 北京大学、武汉大学、中山大学分别有139、83、38人提交。
- 本科生组**182**队，研究生组**107**队。
- 分布于45个一级学科，其中图书馆、情报和档案队伍59组，应用经济学50组，社会学、管理科学与工程、计算机科学、统计学也分别达到27、23、19、19组。
- 参赛队员共提交了约200个数据集至开放研究数据平台。

首届全国高校**数据**驱动
创新研究大赛

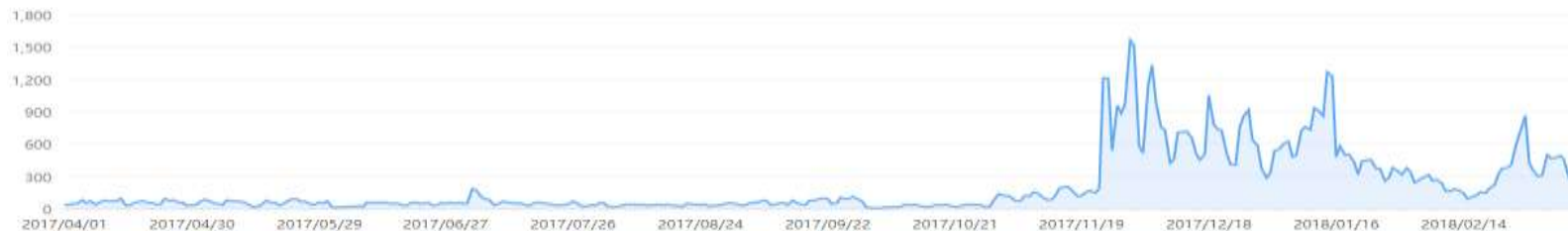
评审结果公布
获奖名单

大赛成果

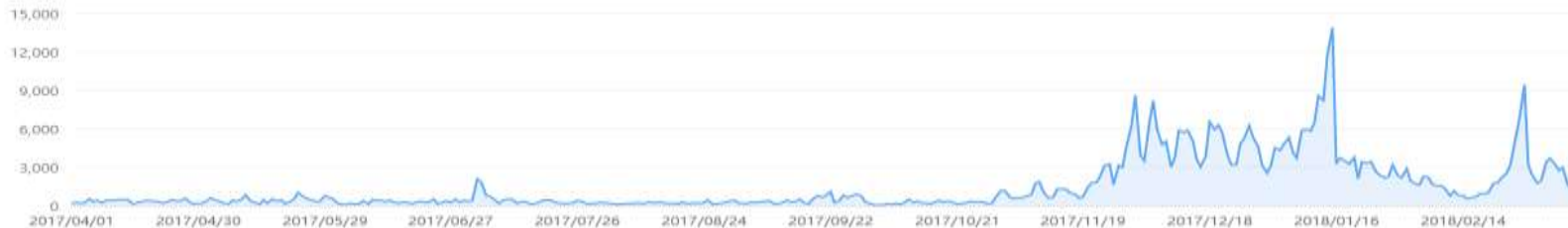
平台流量**十倍**增加

用户数量增加**四倍多**

指标: 访客数(UV)



指标: 浏览量(PV)





北京大学图书馆、北京大学信息管理系统、北京大学中国社会科学调查中心、重庆西部大数据前沿应用研究院联合主办的第二届全国高校数据驱动创新研究大赛（数字中国创新大赛联盟高校专场），截至2019年1月11日，报名工作已全部完成。大赛共吸引来自北京大学、清华大学、复旦大学、浙江大学、中国人民大学、武汉大学、中山大学、北京师范大学、西安交通大学、四川大学、南开大学、吉林大学、重庆大学、重庆邮电大学、河北大学、四川师范大学等199所国内高校的600支队伍、1704名选手参加，覆盖全国29个省、自治区和直辖市，涉及应用经济学、计算机科学与技术、统计学、社会学、图书馆情报与档案管理、管理科学与工程、公共卫生与预防医学等48个学科领域。



数据驱动研究范式和一流高校数据 服务支撑体系研究

——首届全国高校数据驱动创新研究大赛综述

□崔海媛* 罗鹏程 赵静茹 聂华 王继民 张久珍

摘要 数据驱动研究已成为全学科研究范式,数据服务也已成为高校师生最需要的新服务。通过调研国内外数据驱动研究比赛情况,介绍首届全国高校数据驱动创新研究大赛参赛情况,对组织过程、评选方法、论文选题和研究方法进行深入分析。并通过研究各学科数据驱动研究方法的特点和趋势,提出高校数据驱动研究服务支撑体系,为推动数据驱动研究范式和建设高校数据服务体系提供参考。

关键词 数据驱动研究 数据服务 数据管理 数字人文 大数据研究

分类号 G251

DOI 10.16603/j.issn1002-1027.2018.06.005

1 大赛组织

2013年被称为“大数据元年”^[1],2016年被称为“人工智能元年”^[2]。在大数据和人工智能时代,数据的价值受到学术界、产业界和政府的高度重视。为了培养数据人才,很多高校与学会举办了数据大赛,如中国计算机学会的“大数据与计算智能大赛”^[3]、清华大学的“中国高校计算机大赛——大数据挑战赛”等^[4]。为了吸引数据人才解决热点和难点问题,大量竞赛平台在产业界涌现,Kaggle^[5]、DataCastle^[6]、阿里巴巴天池大数据竞赛^[7]等平台吸引了数以万计的数据人才参与。为了推动研究和大数据产业发展,某些省市近年来也举办了一些有影响力的数据比赛,如贵州省的“中国国际大数据挖掘大赛”^[8]、广东省的“广东政务数据创新大赛”^[9]等。在图书情报领域,上海图书馆举办的“开放数据应用开发竞赛”,通过面向全社会征集以开放数据为基础的优秀移动应用产品原型和服务创意,以期充分释放人文数据的价值^[10]。

的环境下,数据驱动研究已经渗透到所有学科,需要面向全学科、研究性的数据比赛来促进高校的教学和科研。与数据受到高度重视相对应,国内研究数据本身的开放程度和管理水平还存在不足,数据管理意识也急需提升。为了促进高校学生基于数据进行研究,提高学术创新能力,促进研究数据的保存、共享和利用,北京大学图书馆、北京大学信息管理学院、南海大数据应用研究院主办,联合国信息中心等多家机构,合作组织了“首届全国高校数据驱动创新研究大赛”^[11](以下简称“大赛”)。大赛一经发布便吸引众多学生参赛报名,在全国高校范围内产生了较大影响。

本文对大赛的组织和参赛情况进行介绍,分析参赛论文选题,以及数据驱动研究方法在各学科的现状与趋势,揭示数据密集型研究范式对各学科的影响;同时,还对数据驱动研究服务支撑体系的设计提出建议,为高校数据驱动研究和服务体系建设提供参考。

三、开设《数字人文》课程 聚焦数据能力培养



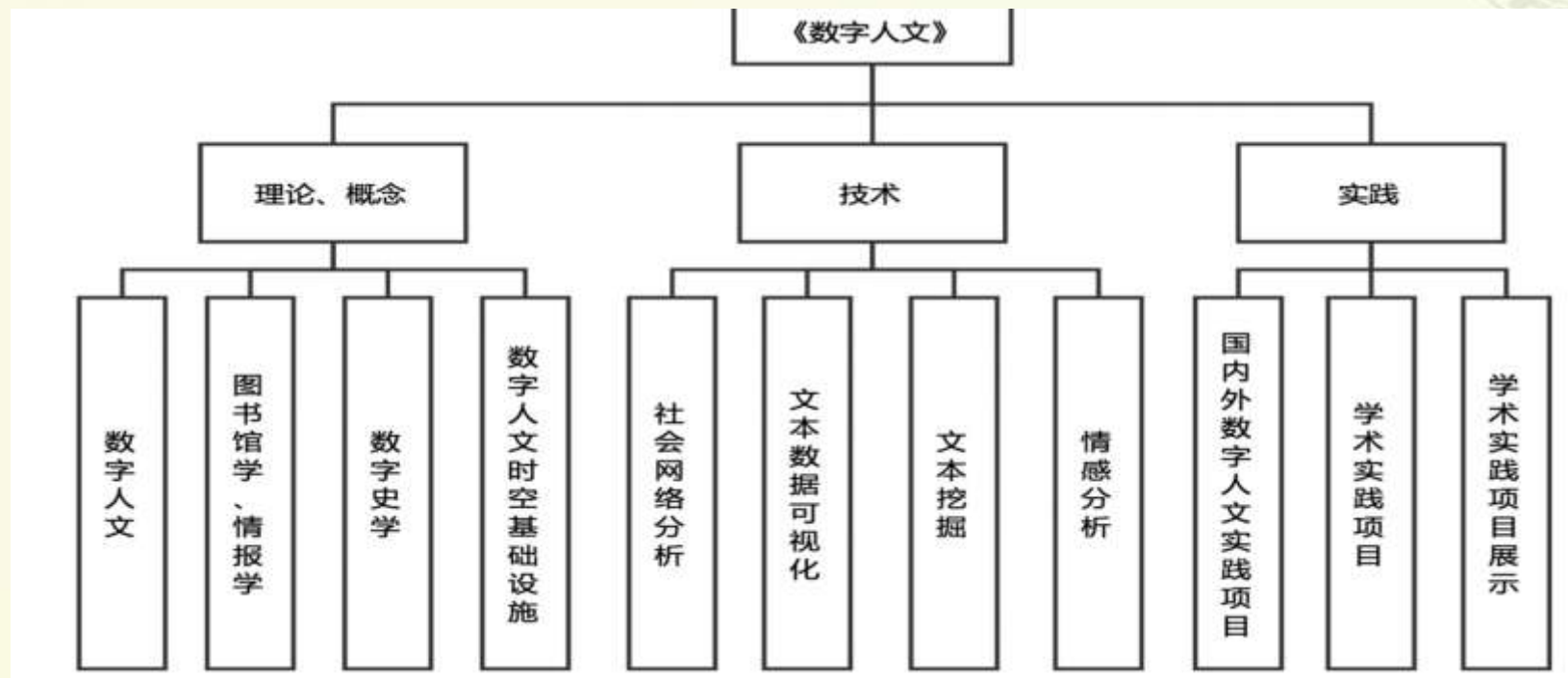
- 我们以数字人文课程为试验点，对本科教育改革和图情专业教育改革进行探索。数字人文基于“知识”这一核心概念，借助数字技术将传统图书馆学对于文献的关注深化到对于文献当中蕴含的知识的挖掘。

课程目标



- (1) 促进图书馆学与情报学课程的融合
- (2) 加强跨学科研究的视角
- (3) 培养学生数字人文项目实践能力

课程内容架构



授课团队

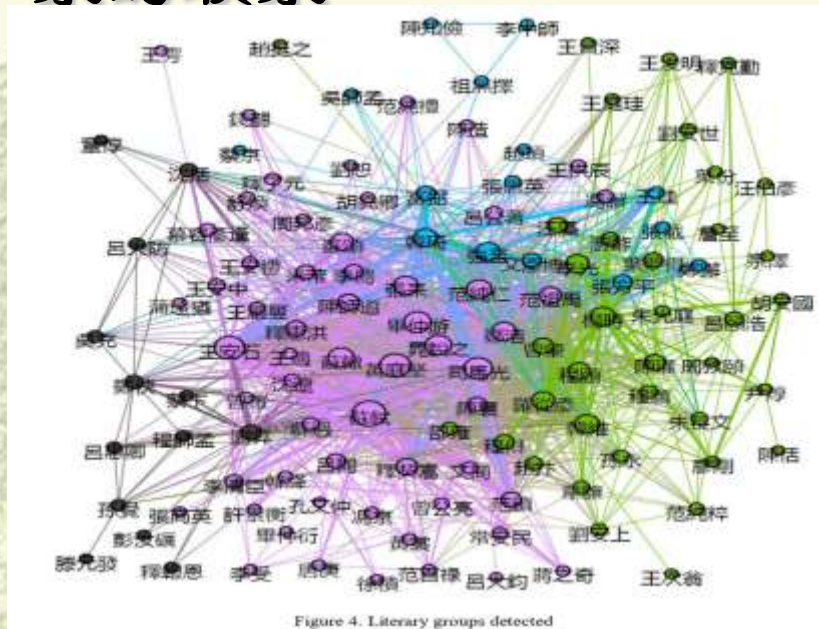


本课程参与授课的教师共 6 人，分别是北京大学信息管理系张久珍教授，其主要研究方向为图书馆学基础理论、参考咨询服务、数字人文；王继民教授，其主要研究方向为搜索引擎、Web 数据挖掘、科学评价学、信息可视化等；张鹏翼副教授，其主要研究方向为信息组织、用户信息行为、人机交互等；黄文彬副教授，其主要研究方向为管理信息系统、多媒体技术等；邀请嘉宾有上海图书馆刘炜副馆长，是图书馆信息技术利用方面的著名学者，同时具有丰富的上海图书馆数字人文项目建设经验；北京大学图书馆朱本军副研究馆员，其具有历史学、教育学背景，是国内较早开展数字人文研究的专家。课程由张久珍教授为总负责人。

数据能力的培养（一）

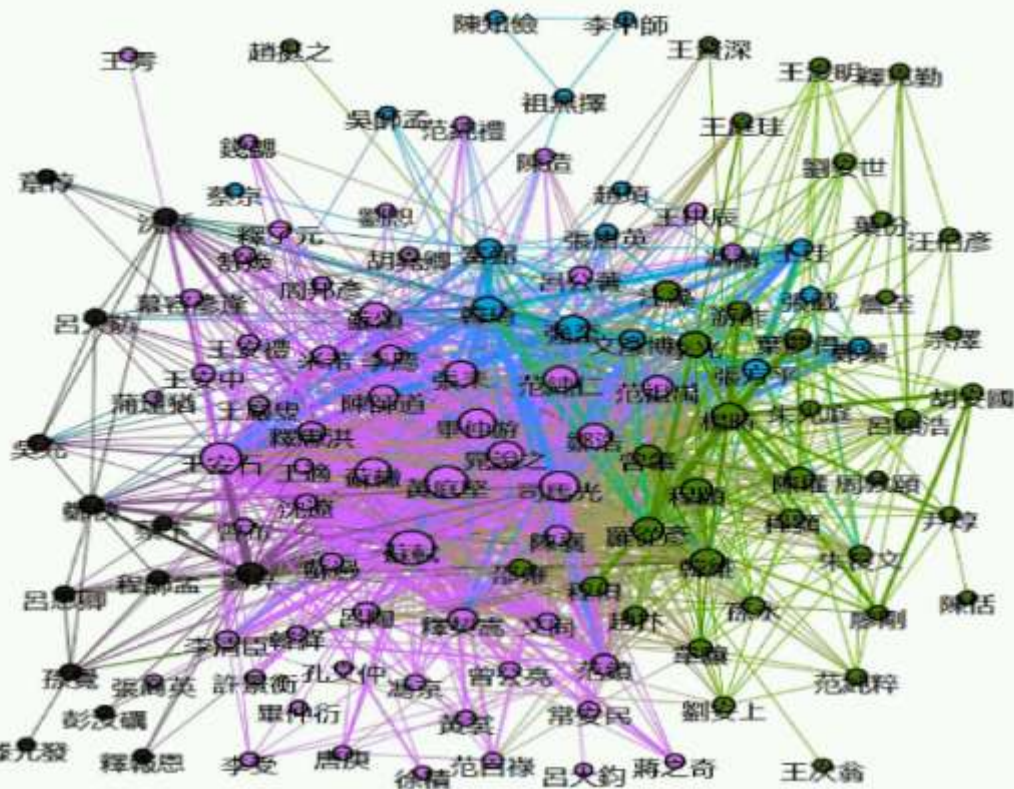


使用CBDB来研究古代中国文学关系和政治关系的联系



2018年3月获哈佛大学Fong Prize
数字人文比赛一等奖，奖金3000美元

来自：北京大学信息管理系学生尚闻一



著述关系矩阵社群发现结果

A紫色：核心群体、蜀学

B绿色：洛学

C黑色：新学

D蓝色：多为仁宗朝老臣

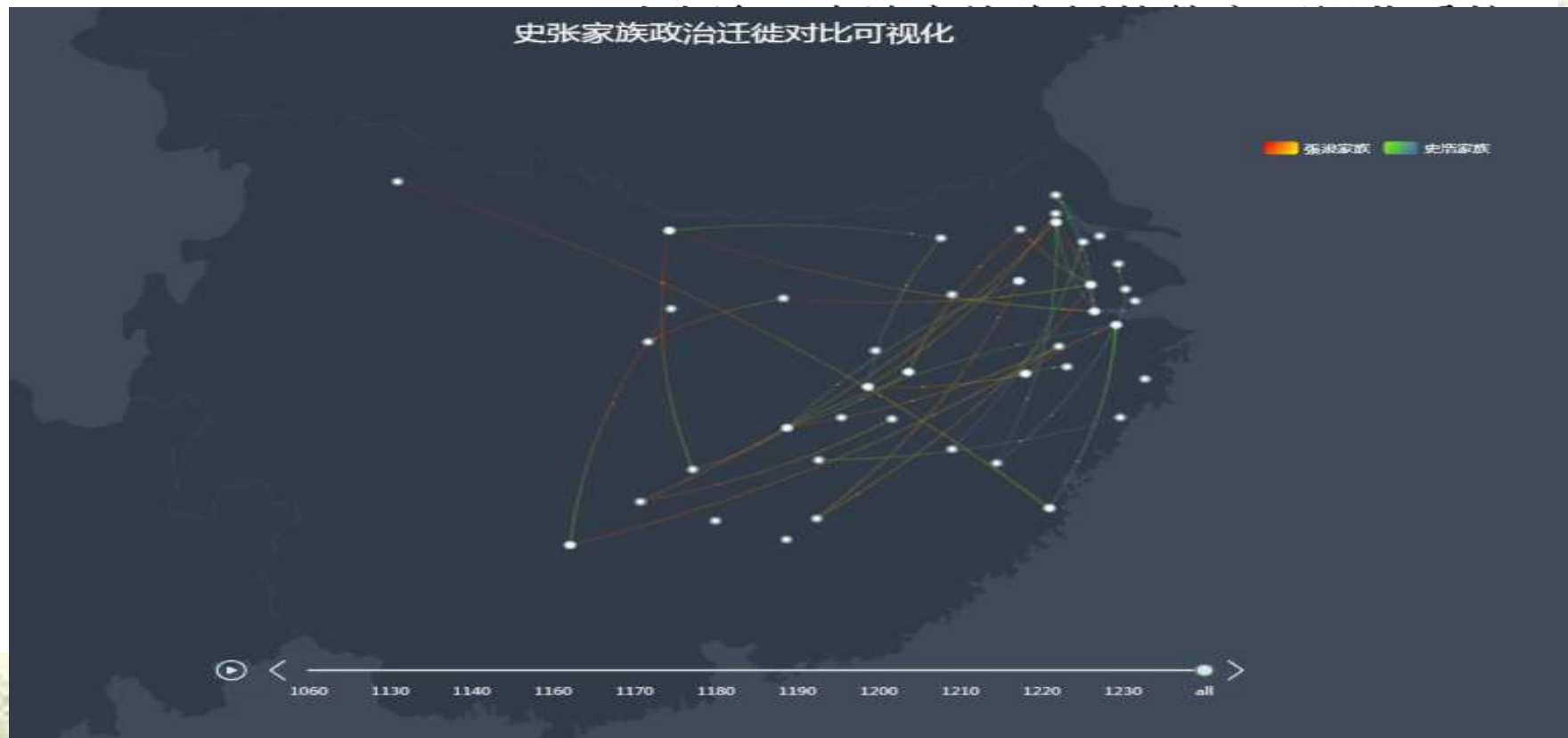
来自：北京大学信息管理系学生尚闻一

信息管理系

北京大学

宋代家族任职比较

史张家族政治迁徙对比可视化



数据能力培养（二）



名称	修改日期
 support documents	2018/11/28 16:09
 民国史料结构化项目汇报_结构化小组	2018/7/4 19:50
 民国史料结构化项目汇报_结构化小组	2018/7/4 19:50

以鲁迅为例的民国人物资料库建设

——民国史料结构化项目汇报

作者：郭晨、李芃蓓、陈祉祁、赵东在¹



目录

一、导言	2
1、项目背景与主要概念	2
2、基本思路	2
二、项目汇报	2
1、时间部分	3
1.1 设计思路	3
1.2 栏位设计	3
2、地点部分	5
2.1 设计思路	5
2.2 行政区划表格	5
2.3 地理情况表格	7
2.4 地理情况表初步成果:	9
3、人物部分	9
3.1 设计思路	9
3.2 栏位设计	10
3.3 民国人物仕宦经历结构化——以鲁迅为例	12
3.4 未来展望	13
4、作品部分	13
4.1 设计思路	13
4.2 栏位设计	13
4.3 民国人物作品结构化——以鲁迅为例	15
三、总结与反思	17
附录:	17
项目分工	17



名称

-  晚清民国人物仕宦经历：以鲁迅为例
-  晚清民国人物作品汇编：以鲁迅为例
-  晚清民国人物作品汇编：以鲁迅为例
-  晚清民国时期地名规范表
-  晚清民国时期时间结构化



ID	国号	帝号	年号	年	干支	生肖	阳历年	阴历年
1	清	宣宗	道光	30年	庚戌	狗	1851年	1950年
2	清	德宗	光绪	7年	辛巳	蛇	1881年	1881年
3	民		民国	19年	庚午	马	1930年	1930年

ID	干支	别名	农历	阳历	闰月
1	乙丑	寐月、桃月、李月、蚕月、桐月、季月、晓春、莺月、樱笋时、桃浪	12	1	不
2	丁酉	壮月、桂月、仲秋、中秋、正秩、仲商	8	9	不
3	戊寅	正月、陬月、孟月、端月、孟春、征月、华月、早春、新正	1	2	不

ID	干支	时	时间段	别称
	子		23-01	夜半
	丑		01-03	鸡鸣
	寅		03-05	平旦
	卯		05-07	日出
	辰		07-09	食时
	巳		09-11	隅中
	午		11-13	日中
	未		13-15	日迭
	申		15-17	晡时
	酉		17-19	日入
	戌		19-21	黄昏
	亥		21-23	人定



第一級行政區劃	非正式行政區劃	第二級行政區劃	第三級行政區劃
省	道	府 直隸州 直隸廳	縣
			散州
			散廳
			縣
直隸區	道	府 直隸廳 直隸州	縣
			散州
			散廳
			縣
		盟	旗
屬國		汗部	旗



第一級行政區劃	非正式行政區劃	第二級行政區劃	第三級行政區劃	第四級行政區劃	
省	府/道/行政督察區	縣	鄉	村/閭	鄰
			山地鄉	村	鄰
			鎮	里/閭	鄰
			縣轄市	里	鄰
		市	區	里	鄰
			原住民區	里	鄰
特別行政區域	道	縣	鄉	村/閭	鄰
			鎮	里/閭	鄰
			縣轄市	里	鄰
	盟部	市	市轄區	保/里	甲/鄰
	基巧	旗	鄉	村/閭	鄰
特別行政地方	盟部	旗	鄉	村/閭	鄰
	基巧	宗			
直轄市/院轄市			區	保/里	甲/鄰



A1	模板码																
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P		
经纬度	地名	别名1	别名2	别名3	别名4	起始年	起始月	起始日	终止年	终止月	终止日	区划类别	经度	纬度	上级区划		
1	安徽省	皖				1644						省	117.0435	30.5112	安徽省	省级	
2	河北省	冀	冀			1928						省	116.3779	39.9306	河北省	省级	
3	遼寧省	遼	奉天省			1929						省	123.4301	41.7869	遼寧省	省级	
4	吉林省	吉				1644						省	126.5767	43.9062	吉林省	省级	
5	黑龍江省					1644						省	123.9751	47.3437	黑龍江省	省级	
6	山東省	魯	東省			1644						省	116.2818	36.1787	山東省	省级	
7	廣東省	粵	廣南	粵東	南粵	1644						省	113.2643	23.1335	廣東省	省级	
8	江蘇省	蘇	吳			1644						省	118.778	32.0514	江蘇省	省级	
9	山西省	晉	三晉	晉川	山西	1644						省	112.5743	37.8732	山西省	省级	
10	河南省	豫				1644						省	113.6811	34.7701	河南省	省级	
11	順天府					1368			1911			府	116.3796	39.9091	直隸省	府級、州	
12	天津府					1644			1911			府	117.1976	39.1361	直隸省	府級、州	
13	奉天府	瀋陽				1644			1911			府	123.4301	41.7869	奉天省	府級、州	
14	錦州府					1644			1911			府	121.1454	41.1117	奉天省	府級、州	
15	海倫府					1644			1908			府	126.9661	47.461	黑龍江省	府級、州	
16	東昌府	東昌				1368			1911			府	115.9981	36.4469	山東省	府級、州	
17	廣州府	廣東	廣東行省	粵	五羊	1368			1911			府	113.2643	23.1335	廣東省	省级	
18	南京府	金陵	南京	建康	江寧府	1912			1913			府	118.778	32.0514	江蘇省		
19	太原府	并州	朔州	并部	晉陽	618			1911			府	112.5743	37.8732	山西省	省级	
20	开封府	大梁	汴京	南商		907			1911			府	114.3525	34.7847	河南省	省级	
21	北平市	順天府	京兆地方	北平特別市	北京	1911			1949				直轄市	116.3796	39.9091	河北省	1914年
22	南京市	首都特別市	寧	京		1927							直轄市	118.7906	32.0609	江蘇省	1912, 1927
23	大興縣					1368			1958			縣	116.368	39.9314	順天府	縣	
24	宛平縣					1012			1957			縣	116.368	39.9314	順天府	縣	

民国时期行政区划表_地理情况表



	序号	栏目	定义	数据格式
			定义	数据格式
人物基本资料	1	ID	规范的数据识别码	顺序数字
	2	姓	标志家族的字	文字
	3	本姓	原先使用的标志家族的字	文字
	4	名	由人所起的名字，正名	文字
	5	别名/曾用名	正名以外的名字	文字
	6	字	人的表字	文字
	7	别号	名、字以外另起的称号	文字
	8	年号	帝王纪元所立的名号	文字
	9	谥号	古人死后依其生前行迹而为之所立的称号	文字
	10	庙号	皇帝死后，在太庙立室奉祀时特起的名号	文字
	11	赐号	受赐的名号、封号	文字
	12	尊号	尊奉帝后或其先王及宗庙等的称号	文字
	13	小名	乳名，幼时起的非正式的名字	文字
	14	题跋名	在题跋中用于落款的名号	文字
	15	笔名	作者发表作品时所署的别名	文字
	16	化名	为使他人不知真实姓名而改用别的名	文字
	17	法名	皈依佛教后，由法师起的名字	文字
	18	避讳名	因避讳改用的名字	文字
	19	室名	居室、书斋的名号	文字
	20	印章名	常用印章上的名字	文字
	21	性别	生理性别，男性或女性	文字
	22	种族/部族/民族	从属的民族部落等族群共同体	文字
	23	宗教信仰	信仰的宗教及派别	文字
	24	国籍	个人具有的属于某个国家的身份	文字
	25	郡望/籍贯	祖属地方及所属家族	文字
	26	行第	在家族中排行的次序	数字
	27	出生年	出生的年份	公历年，数字
	28	出生月	出生的月份	公历月，数字
	29	出生日	出生的日期	公历日，数字
	30	去世年	去世的年份	公历年，数字
	31	去世月	去世的月份	公历月，数字
	32	去世日	去世的日期	公历日，数字
	33	出生地	出生的地点	权威地名+行政区划级别，文字
	34	死所	去世的地点	权威地名+行政区划级别，文字
	35	葬地	下葬的地点	权威地名+行政区划级别，文字
	36	官职	作为政府官员的职位	政府名称+政府机构+官职，文字
人物仕宦经历	37	任职形式	出任某官职的形式，如正式任职、兼职等	文字
	38	入职方式	得到某官位的途径，如科举、恩荫等	文字
	39	起始年	上任的年份	公历年，数字
	40	起始月	上任的月份	公历月，数字
	41	起始日	上任的日期	公历日，数字
	42	截止年	去职的年份	公历年，数字
	43	截止月	去职的月份	公历月，数字
	44	截止日	去职的日期	公历日，数字





23

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	品级	文官阶	文官阶别	武官阶	顶戴	蟒袍	文官补服	武冠补服	文官年俸	文官禄米	文官养廉	武官年俸	武官蔬菜	武官灯红	武官养廉银
1	正一品	光禄大夫		建威将军	红宝石	九蟒五爪	仙鹤	麒麟	180两	180斛	16000两	609两	180两	200两	2000两
2	从一品	荣禄大夫		振威将军	红宝石	九蟒五爪	仙鹤	麒麟	180两	180斛	16000两	609两	180两	200两	2000两
3	正二品	资政大夫		武显将军	珊瑚	九蟒五爪	锦鸡	狮子	155两	155斛	13000两	599两	140两	160两	1500两
4	从二品	通奉大夫		武功将军	珊瑚	九蟒五爪	锦鸡	狮子	155两	155斛	13000两	599两	140两	160两	1500两
5	正三品	通议大夫		武义都尉	蓝宝石	九蟒五爪	孔雀	豹子	130两	130斛	6000两	243两	48两	38两	500两
6	从三品	中议大夫		武翼都尉	蓝宝石	九蟒五爪	孔雀	豹子	130两	130斛	6000两	243两	48两	38两	500两
7	正四品	中宪大夫		昭武都尉	青金石	八蟒五爪	雁	虎	105两	105斛	3700两	141两	18两	24两	260两
8	从四品	朝议大夫		宣武都尉	青金石	八蟒五爪	雁	虎	105两	105斛	3700两	141两	18两	24两	260两
9	正五品	奉政大夫		武德骑尉	水晶	八蟒五爪	白鹇	熊	80两	80斛	2400两	90两	12两	12两	200两
10	从五品	奉直大夫		武德佐骑尉	水晶	八蟒五爪	白鹇	熊	80两	80斛	2400两	90两	12两	12两	200两
11	正六品	承德郎	宣德郎	武略骑尉	砗磲	八蟒五爪	鸳鸯	彪	60两	60斛	1250两	49两			120两
12	从六品	儒林郎		武略佐骑尉	砗磲	八蟒五爪	鸳鸯	彪	60两	60斛	1250两	49两			120两
13	正七品	文林郎		武信骑尉	素金	五蟒四爪	鹞鹞	犀牛	45两	45斛	1200两	36两			90两
14	从七品	征仕郎		武信佐骑尉	素金	五蟒四爪	鹞鹞	犀牛	45两	45斛	1200两	36两			90两
15	正八品	修职郎		奋武校尉	阴文镂金	五蟒四爪	鹞鹞	犀牛	40两	40斛	40两				
16	正九品	修职佐郎		奋武佐校尉	阴文镂金	五蟒四爪	鹞鹞	犀牛	40两	40斛	40两				
17	正九品	登仕郎		修武校尉	阳文镂金	五蟒四爪	练雀	海马	33两1.14	33斛1.14	33两1.14钱				
18	从九品	登仕佐郎		修武佐校尉	阳文镂金	五蟒四爪	练雀	海马	33两1.14	33斛1.14	33两1.14钱				
19	未入流					五蟒四爪	黄鹞	无	31两5钱	31斛5斗					
20	一级														
21	二级														
22	三级														
23	四级														
24	五级														

主表: 人物基本信息, 人物仕宦经历, 职官表, 官员品级, 官员任职形式, 官员入职方式, 事件

信息管理系

北京大学



ID	姓	名	曾用名	字	小名	笔名1	法名	性别	种族/部族/民族	宗教信仰	出生年	出生月	出生日
1	周	树人	樟寿	豫才	阿张	鲁迅	长庚	男	汉	无	1881	9	25

ID	官职	任职形式	入职方式	起始年	起始月	起始日	截止年	截止日	截止时
1	南京临时政府教育部社会教育司第一司	正式任职	委任	1912	3		1912	5	5
2	北洋政府教育部社会教育司第一司	正式任职	委任	1912	5	5	1912	8	
3	北洋政府教育部金事	正式任职	荐任	1912	8		1917	7	7
4	北洋政府教育部金事	正式任职	荐任	1917	7	14	1925	11	

去世年	去世月	去世日	国籍	郡望/籍贯	行第	出生地	死所	葬地
1936	10	19	中国	浙江省绍兴府	1	浙江省绍兴府会稽县东昌坊口新台门	上海特别市	上海特别市虹桥万国公墓

Figure 15: 表集，鲁迅基本资料与仕宦经历表

B	C	D
	栏位	注释
基本时间背景	写作开始时间	指根据年谱, 该本作品该人物开始动笔的日期, 以公元纪年表示, 格式为年-月-日, 如(1900-01-01, 月或日不详者
	成文(成书)时间	指根据年谱, 该本作品该人物结文的日期)
	成文(书)作者时年	指该人物或作者完成该作品时的年龄(按虚岁计算, 如示例中鲁迅以《鲁迅年谱》-黑龙江人民出版社1979年3月出版
	政府背景	根据该人物生卒年份, 为清政府/民国政府, 即表示著作当年的当政政府)
基本信息	作品名称	即该人物在该时段内(或某年某月某日)所著文章或书籍
	作品别名或曾用名	指该作品的其他名称, 或曾经使用的名称
	体裁	指该篇作品的文学样式和类别, 如随笔、散文、小说、杂文、日记等等
	署名	指该人物发表该作品时所署名, 无记录则保持空白
	合作作者	指与该人物合作完成该作品的人, 没有或无记录可查则保持空白
	属性1	用于区分该作品是否是译本, 如是, 则填“译本”, 不是则保持空白
	属性2	用于区分该作品是否为单行本, 如是, 则填“单行本”, 不是则保持空白
	地位/荣誉	指当时或后世社会或他人给予该作品的评价及其获得的文学或社会地位, 和其所获得的荣誉等
	备注	所列各项未能叙述完整者, 于此处做简要补充说明
	译本选填	原著作者
原著作者国籍		指该人物翻译的文章或单行本的原著作者国籍,
转译作者		指该人物翻译的文章或单行本转译自谁何人, 填写该作者姓名, 如多人可增加字段
转译作者国籍		指该人物翻译的文章或单行本转移作者的国籍
刊录信息		指该作品发表所在的刊物或作品(作品集)名称, /前为卷数、后为期号数, 无/则无期号数, 如《新青年5/2》, 即作第5卷第2期;
	刊载地	若为他人作品, 格式为《作品名》-作者名
	刊载时间	指该作品被发表在上述刊物或作品的时间
	收录文集	指收录该作品的文集或作品集
	收录时间	指该作品被收入对应文集的时间
	作品调整	指出版前对作品内容、收录情况等各方面所作的调整, 可添加多个字段, 基本格式为“增补/删除-修改内容”
	调整时间	指进行上述调整的时间



数据能力培养（三）

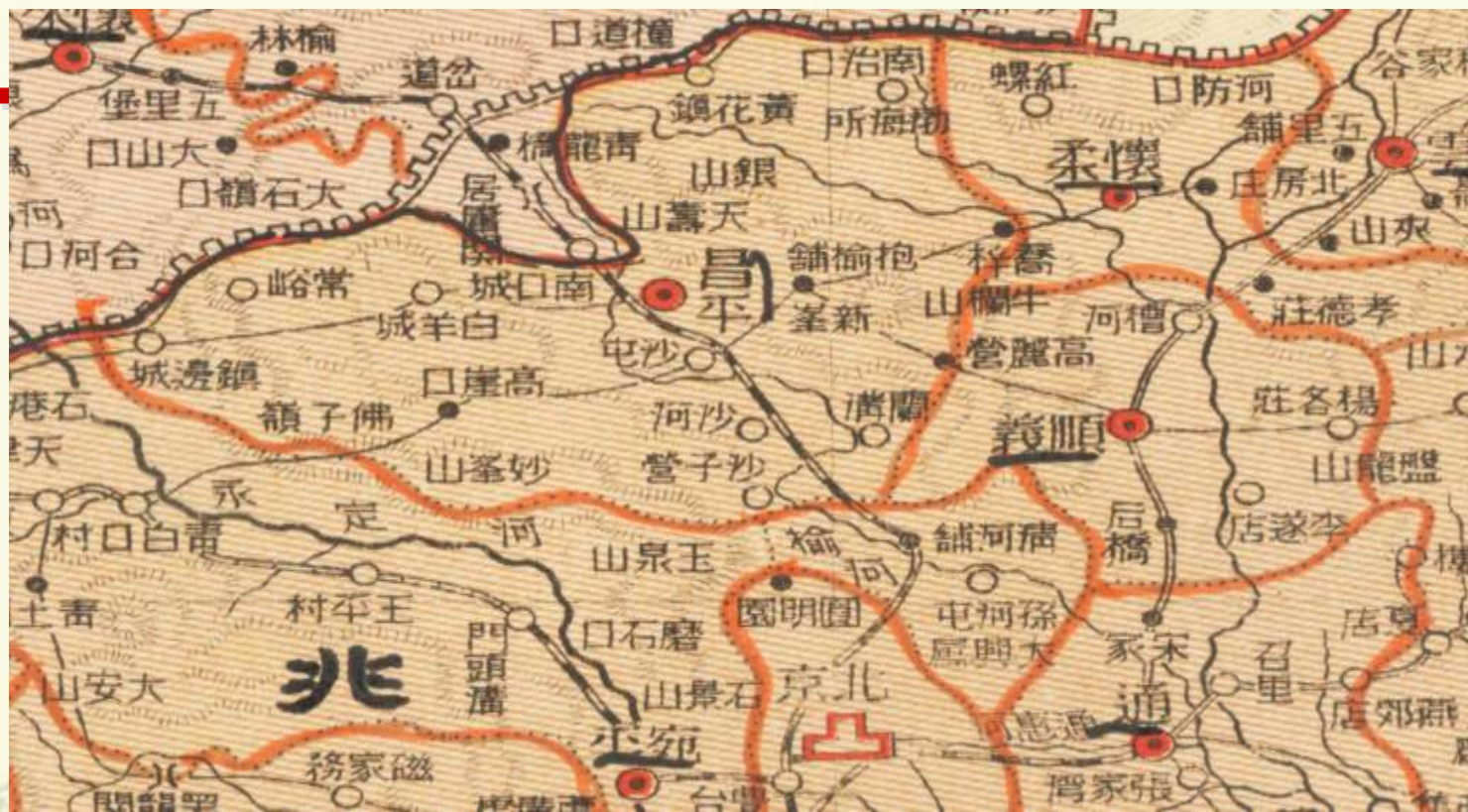
□ 非结构化数据→结构化数据
数据意识与数据能力训练

1. 学会读民国地图及要素
2. 每人选取1张分幅地图，尝试将地图上的非结构化数据转成结构化数据，并录入系统



圖例
(分圖倣此)

大路	驛路	電線	鐵道	未成鐵道	已成鐵道	縣界	道界	省界	國界	水道	山脈
隄防	驛站	砲台	關隘	商港	市集	要邑	首府	屬部	縣治	道治	省治





中华民国地理规范库（民国四年|1915年）著录系统

1. 漢語地名 * 别名1 别名2 英語名稱 *

2. 起始紀年方式 = 民國國號紀年 = 民國年號 = 民國年
 國號紀年 中華民國 民國 四年

3. 起始紀月日方式 是否閏月 農曆月份 十一月農曆/月相日
 農曆紀月 閏 十一月 请选择

4. 起始方式
 请选择

5. 是否为政区
☒ 是 ☐ 否

6. 它属于
☒ 要邑 ☐ 市集 ☐ 舊治 ☐ 遷治 ☐ 縣治
☐ 屬部會府 ☐ 道路 ☐ 驛路 ☐ 大路 ☐ 商港
☐ 山岳 ☐ 河流 ☐ 湖泊 ☐ 沙灘 ☐ 关隘
☐ 飛行場 ☐ 無線電信局 ☐ 寺廟 ☐ 井泉 ☐ 墓碑
☐ 名勝旧迹 ☐ 矿产地 ☐ 已成鐵路 ☐ 未成鐵路 ☐ 炮台
☐ 驛站 ☐ 堤防 ☐ 長城 ☐

8. 結束紀年方式 = 民國國號紀年 = 民國年號 = 民國年
 國號紀年 中華民國 民國 四年

黃花
鎮

Huanghua
zhen



已录入的图幅

- ☐ 京兆
- ☐ 直隶
- ☐ 辽宁
- ☐ 吉林
- ☐ 黑龙江
- ☐ 河南
- ☐ 山东
- ☐ 江苏
- ☐ 安徽
- ☐ 江西



序号	地名	地名拼音	属于	政区隶属于	属于哪个省	属于哪个道	属于哪个县	经度_X	纬度_Y	起始纪年方式	起始年	起始月_农历	动作
50	扶餘縣	Fuyu Xian	县治		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
51	伯都訥站	Bodune Zhan	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
52	社里	Sheli	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
53	長春嶺	Changchunling	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
54	浩包	Haobao	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
55	新發屯	Xinfa Tuntun	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
56	小城子	Xiaochengzi	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
57	石頭城	Shitou Cheng	市集		吉林省	滨江道	扶余縣			中华民国	四年	十一月	新置
58	榆樹縣	Yushu Xian	县治		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
59	牛頭山	Niutou Shan	山岳		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
60	大嶺	Daling	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
61	孤榆樹	Guyu Shu	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
62	西河城	Xihe Cheng	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
63	五棵樹	Wuke Shu	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
64	秀水甸	Xiushui Duan	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
65	土桥子	Tuqiao Zi	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
66	大新立屯	Daxinli Tuntun	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
67	黑林子	Heilin Zi	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
68	法特哈	Fate Ha	市集		吉林省	滨江道	榆樹縣			中华民国	四年	十一月	新置
69	雙城縣	Shuangcheng Xian	县治		吉林省	滨江道	雙城縣			中华民国	四年	十一月	新置
70	報馬川	Baoma Chuan	市集		吉林省	滨江道	雙城縣			中华民国	四年	十一月	新置
71	車家崴	Chejia Wei	市集		吉林省	滨江道	雙城縣			中华民国	四年	十一月	新置
72	朱仙鎮	Zhuxian Zhen	要邑		河南省	开封道	開封縣			中华民国	四年	十一月	新置
73	仙人庄	Xianren Zhuang	市集		河南省	开封道	開封縣			中华民国	四年	十一月	新置
74	仇店	ChouDian	市集		河南省	开封道	開封縣			中华民国	四年	十一月	新置
75	蘇村	Su Cun	市集		河南省	开封道	開封縣			中华民国	四年	十一月	新置
76	水窩鋪	Shuiwo Pu	市集		河南省	开封道	開封縣			中华民国	四年	十一月	新置

4100个民国四年数据

北京大学“数字人文”课程教学实践及经验探索

■ 张久珍 韩豫哲

北京大学信息管理系 北京 100871

摘要: [目的/意义]在新时代高校本科教育改革的背景下,北京大学信息管理系以数字人文教育为试验点,开设“数字人文”课程,为数字人文课程教学提供参考,对数字人文教育进行探索。[方法/过程]介绍并总结该课程的培养目标、内容设置和特色等,以此探讨数字人文课程建设对于图书馆学、情报学学科,以及本科教育改革的示范性作用。[结果/结论]对数字人文课程建设提出建议,包括组建教师团队、打造应用场景和将数字人文课程作为基础素质教育课程等。

关键词: 本科教育改革 数字人文 数字人文教育 图书情报学教育

分类号: G250 G642

DOI: 10.13266/j.issn.0252-3116.2019.19.001

1 引言

信息时代的变革强烈地冲击和影响图书馆学专业教育。iSchools 联盟重视人与信息、技术三者之间的融合,为图情教育提供了新的发展理念,也优化了图情职业环境。国内图情学科教育改革呈现了以信息为核心、重视技术应用、重视社会需求、面向更广泛的信息职业、培育复合型人才等特点。李杨^[1]通过对北京大学信息管理系招生、培养和就业情况及国内外图书馆情报学重点高校的人才培养模式的分析,提出了信息技术在图书馆学情报学人才培养方面日益重要、图书馆学情报学需要进行协调整合等趋势。

2016年,《北京大学本科教育综合改革指导意见》

在课程开设准备阶段,课程筹备组对伦敦大学学院、印第安纳大学等大学的数字人文课程体系进行了调研。而国内的数字人文教育刚刚发端,王涛^[1]介绍了他面向历史研究开设的“数字工具与世界史研究”课程,进行了经验总结和反思。笔者将以北京大学信息管理系“数字人文”课程为例总结在本科教育改革背景下的数字人文课程建设的经验。

2 北京大学信息管理系“数字人文”课程介绍

2.1 课程目标

2.1.1 促进图书馆学与情报学课程融合 图书馆学



结束语



- 国内高校及高校图书馆在数据素养教育上已经取得了一定的探索成果，但当前国内的数据素养教育还未广泛展开，合作机制有待建立，课程体系有待完善，教育内容有待丰富，教学方式有待创新，相关的开放数据平台等数据基础设施建设也有着较大的发展空间。



结束语

- 在高校师生、科研人员、工作人员以及社会公众对于数据素养教育旺盛的需求下，数据素养教育将成为信息素养教育的重要内容和发展方向。

A scenic view of a lake, likely Xihai Lake at Peking University, with a traditional Chinese pagoda on the left and lush green trees along the shore. The sky is blue with scattered white clouds. The text "感谢观看!" is centered over the image.

感谢观看!