

北京邮电大学

大学生创新训练项目 立项申请书

项目来源：导师科研类☒ 自主探索类☐

校企合作类☐ 滚动支持类☐

科研院所合作类☐ 校际合作类☐

雏雁获奖类☐

项目名称：_____跨学科合作网络发现与推荐_____

项目名称(英文)：_____Interdisciplinary Collaboration Network
Discovery and Recommendations_____

项目依托学院：_____计算机学院（国家示范性软件学院）_____

项目负责人：_____胡逸同_____

联系电话：_____13661192968_____

E-mail：_____wiederholung@outlook.com_____

指导教师：_____王祎_____

E-mail：_____yiwang@bupt.edu.cn_____

起止年月：_____05/2022-04/2023_____

填报时间： 2022 年 4 月 26 日

填写说明

- 1、 本申请书所列各项内容均须实事求是，认真填写，表达明确严谨，简明扼要。
- 2、 申请书为大 16 开本（A4），在网上下载后，根据填报项目类别可自行删减和加页，但须保持格式和内容与原件一致。
- 3、 第五部分“推荐、评价及审批意见”不用填写。

一、基本情况

项目名称	跨学科合作网络发现与推荐						
项目负责人	胡逸同	学号	2020213350	所在学院	国际学院	手机号	13661192968
		专业	电子商务及法律	班级	2020215111	邮箱	wiederholung@outlook.com
指导教师	王祎	职称	教授	所在学院	计算机学院 (国家示范性软件学院)	手机号	13127958790
					邮箱	yiwang@bupt.edu.cn	
项目类别 (类别说明见立项指南)	<input type="checkbox"/> 华为众智 <input type="checkbox"/> 小米专项 <input type="checkbox"/> 智能硬件 <input type="checkbox"/> 社交媒体 <input type="checkbox"/> 数字娱乐 <input type="checkbox"/> 通信网络 <input type="checkbox"/> 医疗健康 <input type="checkbox"/> 公共服务 <input type="checkbox"/> 电子商务 <input checked="" type="checkbox"/> 教育文化 <input type="checkbox"/> 房产家居 <input type="checkbox"/> 理论研究 <input type="checkbox"/> 机器人 <input type="checkbox"/> 无人机 <input type="checkbox"/> 智能制造 <input type="checkbox"/> 智能交通 <input type="checkbox"/> 创意设计 <input type="checkbox"/> 其他_____						
检索关键词	推荐系统, 交叉学科, 网络图谱, 语义分析, 网络可视化						
项目成员 基本信息	姓名	学院	专业	班级	学号	电话	邮箱
	王伊哲	国际学院	电子商务及法律	2020215111	2020213355	15353519783	anonymity@bupt.edu.cn
	胡逸同	国际学院	电子商务及法律	2020215111	2020213350	13661192968	wiederholung@outlook.com
	巢环宇	国际学院	电子商务及法律	2020215111	2020213349	15580092683	chaohuanyu@bupt.edu.cn
	朱子炫	国际学院	电子商务及法律	2020215111	2020213363	13935559636	ternura@bupt.edu.cn
	王俊翔	国际学院	电子商务及法律	2020215111	2020213354	18910761281	1034696975@qq.com
团队主要成员介绍	<p>王伊哲: 编程是爱好, 擅长 C/C++, Java, Python, JavaScript, XML, CSS 以及 NoSQL 和 SQL 的应用, 正在学习 Linux 操作系统及开发。在中学曾参加过算法比赛, 大学后参加雏燕、挑战杯等活动, 独立开发过微信程序, 在本次项目中主要负责后端源码等工作。</p> <p>胡逸同: 高中获有全国青少年科创大赛二等奖。学习速度快, 动手实践能力强。熟悉计算机领域的基础概念, 掌握 C, Python, Java, SQL 基础, 对计算机网络和 Linux 有较多的探索。在本次项目中负责推荐算法和数据库。</p> <p>王俊翔: 擅长 C/C++, java, python, MySQL 等等技术, 正在学习网页前后端技术。高中荣获市级创客马拉松一等奖。大学参与雏燕与挑战杯。在本次项目中负责前端开发。</p> <p>朱子炫: 熟悉计算机领域基础概念, 掌握 C/C++, Java, Python 等编程语言以及数据库的基本操作, 使用过 Linux 系统, 在本次项目中负责数据库和后端。</p> <p>巢环宇: 学习过 C/C++, Python 编程语言, HTML5, CSS 等前端技术, 会使用 MongoDB</p>						

	数据库语言，在本项目中主要负责数据库和前端。
指导教师承担 科研课题情况	目前，指导教师承担两项国家级项目：（1）面向人与社会因素的软件工程（国家海外高层次人才计划青年项目），2022-2024，主持；（2）基于博弈网络和社会信号理论的软件开发团队合作理论与技术研究（国家自然科学基金面上项目），2022-2025，王祎，主持
指导教师对本 项目支持情况	本项目是指导教师与我校人事处合作的示范应用。在本项目中，指导教师将在数据收集和清洗，网络图谱构建，推荐系统实现，可视化等技术领域进行指导。项目参与学生可使用指导教师目前的软硬件资源，包括 AMD ASPY7432+NVida3090*2 高性能计算和深度学习平台。指导教师的部分科研经费也可供项目参与学生使用。

二、立项依据

(1) 项目创意来历及项目意义

跨学科学术合作已经成为学术创新的重要机遇。然而，发现潜在的学术合作机会对于研究者个人而言仍较为困难。本项目旨在通过对学者的研究成果进行语义分析，构建基于语义、社会关系等因素的多维合作网络，从而实现对跨学科潜在学术合作关系发现与推荐。我校虽然目前学术实力和科研水平都极为雄厚，但是如果在此基础上，可以通过这个平台来强化不同院系导师之间的交叉合作，将更有利于发掘出潜在的科研方向与机遇，应用在我校，将推动校内科研合作的进一步发展。

与之同类型的有清华大学和以色列特拉维夫大学共同构建的 XIN 中心，XIN 中心的宗旨，在于利用两所高校雄厚的科研实力及双方在各自国家教育领域的领军地位，努力推动学术进步和跨学科科研创新。所以我们的项目与之不同的地方主要在于，我们更倾向于发掘校内的潜在学术合作关系，这将有利于我们更好地发现校内科研合作的优势与痛点。

在该平台，用户可以上传个人研究汇总文件，包括个人基本信息，个人发表文献数，合作导师信息，合作导师发表文献数等。系统可以解析上传的文件，判断信息符合要求后存入数据库，并作出响应。用户也可以直接在平台上执行查询功能，查询本人或其他导师的相应科研合作图，图将以网状结构呈现现有的以及潜在的合作关系。

(2) 项目研究主要内容

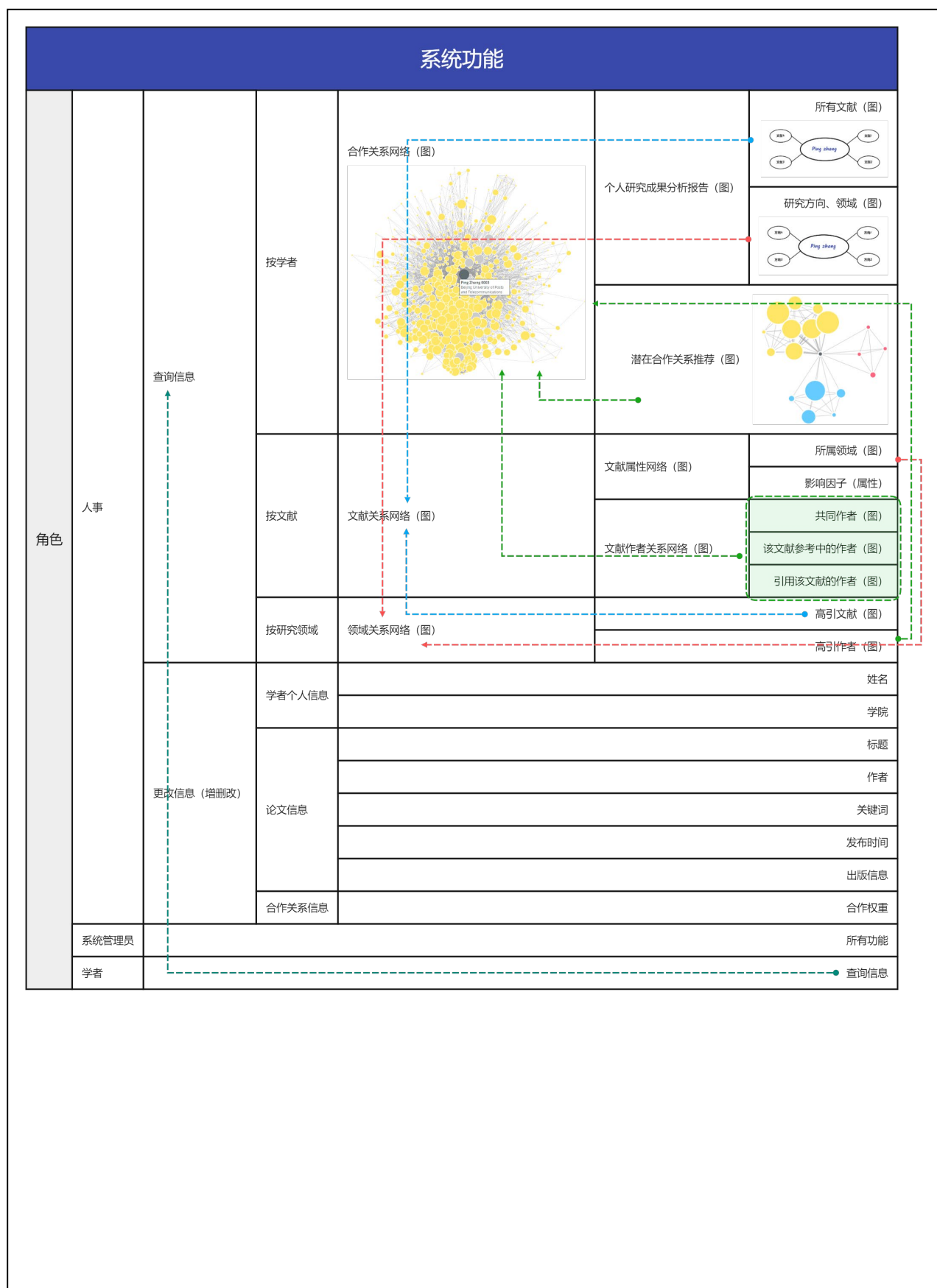
项目主要内容包括但不限于对我校师生的科研学术论文进行爬取，构建学术合作数据库，在数据库中分析和梳理其中的跨学科学术合作关系网络，并以此为基础对其进一步进行分析和学习推理，给出用户潜在的科研学术合作伙伴推荐。相较于传统的跨学科学术合作，我们的创新点或者优点就是可以更快速、更广泛的搜索潜在的跨学科的科研合作伙伴，并优化给出合理结果，并面向我校作为学科特色化高校的特殊需求。在后期，我们会开发相关的用户界面和可视化方案，例如通过形象灵动的网络可视化，动态推荐，智能交互等工具，将个体间合作关系、推荐的潜在合作关系等内容形象地展示。

该系统由四大模块构成，分别是数据管理，推荐系统模块，网络可视化和 Web 交互。

其中，数据管理模块负责管理学者的信息，包括收集学者个人信息、发布的文章、合作学者等，并支持增删改查的数据库功能。

推荐系统模块内含统计函数、语义分析、评估函数和推荐函数四大主要部分。该系统首先统计某一学者的发刊数量、合作学者数量、文章关键词以及影响因子等表层数据，作为对该学者评估的基础指标。再对其文章进行语义分析，深度挖掘学者画像，进而由评估函数生成该学者在研究领域多样性，垂直领域研究深度等多个维度的评分，最后由推荐函数推荐该学者在各个领域契合度高的可合作学者，即跨学科潜在合作关系推荐。

网络可视化模块实现了关系/属性网络两大渲染引擎。关系网络是以特定实体为中心，其他实体作为节点与之相连，由同一类型实体（学者 or 文献 or 领域）构成的关系图，用于可视化同一实体间的联系；同理，属性网络是以特定实体为中心，其各个属性的值作为节点与之相连，隶属于同一属性的节点聚成一类，构成的关系图，是关系型数据库的可视化表达，聚焦于特定实体的属性。渲染引擎负责根据前端返回的参数，生成相应的关系图，并将代码片段嵌入 HTML 中。Web 交互模块用于信息展示和用户交互，由账户系统，信息查询，信息更改，网络交互等主要页面构成。账户系统可确认用户角色，包括人事和学者；用户可通过检索特定学者、文献或研究领域查看关系/属性网络；学者只能向人事发起信息变更申请，人事有权对数据库增删改查；网络交互指，网络中每一节点，都具有响应鼠标悬停显示信息、鼠标拖拽自组织动效和鼠标左击进入该节点关系/属性网络等事件的功能。



（3） 项目创新点与项目特色

本项目基本功能为将学者间合作关系进行直观的网络化展示并通过算法智能推荐潜在合作对象、合作领域与合作内容。本项目响应学术界号召，为学者们提供一个可以直观反应本人跨学科合作程度的平台并督促广大学者加强与他人协作并通过跨学科合作产出高价值内容。大多数情况下人们更重视学者本身的科研成果数量但是并不重视或没有途径来考量学者跨学科能力，本项目通过统计与计算将其可视化输出以共学者们考量与自我改进。本项目创新的采用网络化输出将先前只用作记录的合作信息转变为有价值的信息加强学者间的合作与沟通。

（4） 系统方案和技术路线

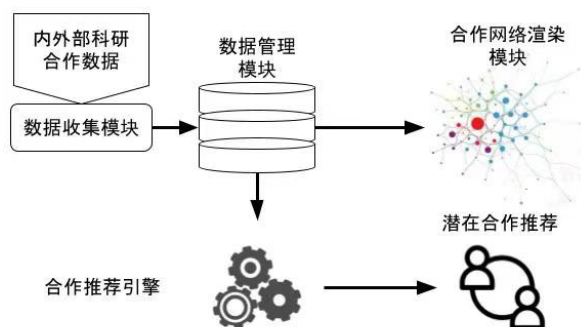
轻量级的前端框架：项目通过轻量级的 Flask 架构网络接口和服务器，较为简易的搭建框架网页，方便前期搭建与成本后期对页面的维护。

基于 Python 的后端技术架构：本项目关键技术为 python 提供的 Crawler，通过爬取 dblp，对学者的研究成果进行语义分析，构建基于语义、社会关系等因素的多维合作网络，从而实现对跨学科潜在学术合作关系发现与推荐。

使用 NoSQL 的数据存储方式：本项目采用 NoSQL 存储和添加数据，相比于正常的关系型数据库，方便建立集群联系和后期对数据库的维护和扩展。此外，MongoDB 支持远程端口链接，方便不同开发者进行数据同步。

图形可视化分析技术：可视化要首选通过分析用户的需求，晚上系统模块功能，然后进行 GUI 设计，通过需求分析可确定未来产品的设计方向和设计理念。确定明确方向后开始进行代码变现。

1. 系统模块



系统模块		
Web	账户系统	
	信息查询	
	信息更改	
	关系/属性网络交互	
网络可视化	关系网络渲染（学者/文献/领域）	
	属性网络渲染（学者/文献/领域）	
推荐系统	统计函数	
	语义分析	
	评估函数	
	推荐函数	
数据管理	收集	爬虫
		API
		Excel
	删除	
	更改	
	查询	

结合实际应用与操作需求，系统将主要分为数据收集模块、数据管理模块、合作推荐引擎和渲染模块，其中最为核心的系统功能模块也将按照功能划分为：数据处理模块、推荐模块和渲染功能模块。系统模块、系统功能子模块设计遵循“高内聚、低耦合”的软件开发原则，将相似的功能内聚为功能模块，模块与模块间建立接口，减少模块间的交互复杂度。以达到系统模块化、功能模块化的目标，从而实现高效实现、逻辑清晰与有效管理的目的。

2. 技术架构

技术栈			
数据	数据获取	Python Crawler	
		dblp search API	
		Excel	
	数据清洗和管理	Python + MongoDB	
	网络图谱	Neo4j	
推荐系统	PyTorch		
可视化与交互	Web	前端	Vue.js
			ECarts
		后端	Flask
Dev	部署	Docker	
	版本控制	Git	
	仓库	GitHub	

本项目涉及 Python Crawler 进行数据获取与挑选，在 dblp 网站上搜索相关老师的论文及个人研究项目；

由 NoSQL 数据库 MongoDB 远程搭建数据库，便于开发成员的链接，存储相关的数据信息；

利用 Python Web 进行网站开发利用到的开发工具有 Vue 与 Flask 分别进行网页的架构和服务器的搭建；用 ECharts 进行可视化和页面美化工作。

3. 技术亮点

本项目基于实际需要通过多种 Web 开发、网络分析、人工智能技术，构建系统，具有一下技术亮点：

1. 满足我校科研管理和进一步发展的实际需要；
2. 轻量化，信息可实现实时更新，可部署于异构的环境中
3. 充分考虑了与现有信息门户、教师画像等应用的整合
4. 具有丰富的交互方式
5. 系统使用效果可量化评估

4. 开发环境

基于 Windows 系统下的开发项目，主要使用开发软件如下

开发语言：Python, Java, JavaScript

集成开发工具：VsCode, VisualStudio, PyCharm, sublime text, datagrip

开发框架：Vue, flask, React.js, ECharts

数据库：MongoDB

第三方接口：阿里云等

项目管理工具：Git, GitHub

(5) 项目进度安排

项目前期大创小组成员每周花费七个小时左右的时间自主学习 python、数据库、前端等等知识并自主熟悉与本项目相关的技术。本项目于 2022 年 1 月 1 日正式完成前期学习进入中期建设。

项目中期主要分为获取整合数据、数据库搭建、前端网页及相关功能搭建、前后端连接等等几个部分。目前数据的获取与整合已经完成，小组正在全力进行数据库的搭建与前端网页搭建的准备。本项目预计于 2022 年 6 月可以完成基础测试版，进行功能测试。

项目后期主要工作为进一步完善项目功能排除漏洞，并进一步扩大项目涵盖学者范围。待基本功能完善之后本项目将着手准备通过算法为学者智能推荐潜在合作对象、合作领域与合作内容板块的搭建与测试。后期预计于 2022 年 8 月开始。

(6) 已有基础

本项目小组成员已经初步完成所需知识的学习与熟悉。

项目初期所需数据已经通过 dblp 学术网站进行获取与整理。

目前本项目主要为电脑代码开发，研究条件齐全，项目正在稳步进行中。

(7) 预期成果形式，达到的技术指标

主要预期成果有如下形式：

- a) 软件应用系统一套，实现系统的主要功能，在校内进行部署，在此基础上形成相应的软件著作权；
- b) 在合作分析、可视化、及推荐等关键方面，形成专利；
- c) 发表相关高水平论文 2-3 篇。

三、经费概算（单位：元）


（一）项目总经费：25774.18
（二）经费分项预算：需附《北京邮电大学大学生创新训练项目经费预算表》

四、成员分工

姓名	项目前期调研学习情况	任务分工	预期投入精力	签字
胡逸同	已经局部完成	后端+数据分析	每周 5h 开发	胡逸同
王伊哲	已经局部完成	后端+数据获取	每周 5h 开发	王伊哲
朱子炫	已经局部完成	后端+数据库	每周 5h 开发	朱子炫
王俊翔	已经局部完成	前端	每周 5h 开发	王俊翔
巢环宇	已经局部完成	可视化+前端	每周 5h 开发	巢环宇

注：包括项目负责人

五、推荐、评价及审批意见

指导教师意见与承诺： 该项目对于高校的科研管理工作现代化和进一步的发展具有重要意义，强烈推荐该项目立项，我承诺在项目进行过程中通过多种方式给予充分指导。 <div style="text-align: right;"> 导师签字：  2022 年 4 月 26 日 </div>	
学院推荐意见： <div style="text-align: right;"> 负责人签字： _____ 学院公章 年 月 日 </div>	
专家组评价意见： <div style="text-align: right;"> 组长签字： _____ 年 月 日 </div>	

学校审批意见：

负责人签字：

盖章

年

月

日

附表：

北京邮电大学 大学生创新训练项目经费预算表

项目所在学院：计算机学院（国家示范性软件学院）

日期: 2022.5.13

支出项目类别	支出项目说明	支出金额（元）	测算依据
1、业务费	打印费、复印费、装订费、书费、资料费等费用	2000	例如：书籍、光盘资料的购买
2、仪器设备购置费	购置或试制专用仪器设备,对现有仪器设备进行升级改造等费用	15774.18	域名 metattri.com 61 SSL 证书 2640 阿里云 ECS 285.18 Roam Reach 3400 JetBrains 1700X5 MarginNote3 288 蓝灯 (VPN) 600
3、材料费	芯片、模块、元器件、电路板等低值易耗品费用	0	例如：低值易耗品、报告、论文打印等
4、外协费	支付给外单位的检验、测试、化验、维修、租赁和加工制作等费用	0	例如：产品的加工包装
5、差旅费	开展科学实验（试验）、科学考察、项目调研、学术交流等所发生的外埠差旅费	3000	例如：调研、测试不同环境下系统的拓展应用
6、会议费	学术研讨、咨询、培训等费用	0	例如：学术研讨等
7、专项业务费	版面费、专利申请及其他知识产权事务等费用	5000	例如：论文、版面费、专利申请等

合计（元）	25774.18
-------	----------

项目指导教师签字：王祎

项目负责人签字：胡逸同

项目组其他成员签字：王伊越 朱子炫 巢宇 马俊翔

