方启明 简介

方启明,男,1981 年出生,清华大学博士。现就职于杭州电子科技大学计算机学院,主要研究方向为类脑计算、深度学习、知识图谱等。主持1项国家自然科学基金, 1项浙江省自然 科学基金,发表 SCI 等高质量论文多篇。

王宇翔 简介

毕业于东南大学计算机科学与工程学院，获工学博士学位。现就职于杭州电子科技大学计算机学院。主要研究方向为：**大数据近似查询优化**、**RDF知识库语义查询**、**分布式数据处理**等研究领域。发表SCI期刊、CCF B类会议等高水平论文10余篇。主持 1 项国家自然科学基金，并作为主要成员参与了 1 项国家自然科学基金重大研究计划及 3 项国家基金项目，作为主要成员参与1项浙江省重点研发计划。

本人主要研究方向所涉及到的细分领域包括：近似查询处理、多查询优化、分布式索引、行列混合存储、RDF知识图谱查询、RDF模糊图查询、RDF知识图谱语义表示（涉及人工智能算法）等。

项目介绍

本项目拟面向科技成果转化领域针对科技人才大数据研究智能化语义搜索引擎，通过人工智能、自然语言处理、分布式索引等技术的研究与应用，在传统基于文本匹配的搜索引擎基础上进行语义搜索功能的拓展，提供丰富、灵活的智能化语义搜索引擎。

本项目大致内容概括如下：

1. 拟通过分布式网络爬虫获取全网科技人才及其成果数据，并利用数据清洗、去噪算法实现数据标准化，构建以人才为中心的科技资源知识图谱；
2. 在此基础上利用大数据和人工智能技术实现对知识图谱的理解、训练和学习，并利用solr、spark、hadoop、word2vec等平台、框架、算法等实现对知识图谱的深度使用，包括科技人才及其成果数据的高效搜索和语义化推荐等；
3. 利用知识图谱技术实现科技人才关系挖掘、科研能力分析以及相关领域发展趋势分析等多样化数据分析服务，并利用echarts等可视化工具实现动态数据展示。

本项目将要涉及到的基本技术及工具：

1. 大数据处理技术：基于Hadoop、Spark的分布式爬虫、分布式离线/实时处理技术；
2. 搜索引擎技术：基于luence、solr的搜索引擎部署与拓展
3. 语义理解技术：基于人工智能与自然语言处理的语义化搜索与推荐技术
4. 知识图谱技术：基于neo4j、graphX的知识图谱计算与查询技术。