厦门理工学院研究生科技活动月——成果展示

# 基于微服务架构与知识图谱技术构建无人机专家知识库

**一. 成果完成人及所在单位：陈秋瑾，王晓荣，陈平，聂德伟　　计算机与信息工程学院　指导教师：何原荣**

## 二. 成果简介：

系统设计框架由最初的 SSM 框架集，进一步将系统的框架优化升级整合为目前的微服务架构；对爬取得到的文本数据基础上利用知识图谱技术及机器学习相关算法、echarts 组件、pyecharts 技术等将非结构化文本数据进一步去重、提取、解析、翻译成元数据，最终以结构化数据形式通过知识树、各类统计图表形式（柱状图、折线图、饼图等）、关键字云图等形式可视化地展示在系统界面上；结合数据提取、数据分析算法、知识图谱、知识提取及表示等理论知识，实现系统的自动文摘和语音播放等智能功能，从而将隐形知识转为机构知识，提高知识的利用效率，实现研究所内部的科研工作的协同工作和知识积累。

## 三. 主要研究内容及创新点：

本系统的主要研究模块有以下几部分：**（1）知识获取模块**：采用分布式及增量式相结合的爬虫技术获取互联网海量数据知识，通过数据自动去重、数据结构化提取等清洗技术，自动化归类实现知识元提取；**（2）知识存储模块**：使用 Redis 数据库结合 MySQL 数据库高效且稳定存储数据，在此研究基础上，进一步利用知识图谱技术中的 Neo4j 图数据库，将提取到的元知识通过**“图-节点”**的形式存储；**（3）知识智能管理及表示模块**：建立了包含知识库、搜索引擎、知识激励、知识社区、专家系统、个人知识管理等系统管理模块。通过知识采集、知识树、知识图谱模型和提取技术实现知识的挖掘，达到知识智能利用；**（4）系统架构设计模块**：针对本系统爬取数据的不断增加的现状，系统也对敏捷性、灵活性和可扩展性需求不断增长，对分布式服务器的的数量和请求和处理速度提出了新的要求，因此本系统选择了微服务架构进行开发；（5）**数据可视化分析及知识推理模块**：对爬取到的数据利用 Echarts 组件及 pyecharts 技术进行可视化分析，分析结果以柱状图、折线图、扇形图等形式展示；通过知识图谱、自然语言处理等算法实现分析爬取数据的关键字云图、机构同现及作者同现图。**本系统的主要创新点有以下 3 点：（1）系统开发架构选择**：系统架构由最初的 MVC 框架优化为功能更丰富强大的 SSM 框架集，在此研究基础上，进一步扩展优化系统的功能，将系统单体式架构升级为目前较为流行的微服务架构；**（2）数据提取及知识抽取**：自定义数据提取的规则，利用机器学习算法半自动化提取爬取的文本数据的关键的信息（如文章标题、作者、发布时间、地区等），再利用知识图谱相关算法对有效数据进行知识抽取，从而得到结构化元知识；**（3）数据可视化及知识表示**：对爬取到的数据利用 Echarts 组件及 pyecharts 技术进行可视化分析；通过知识图谱、自然语言处理等算法实现分析爬取数据的关键字云图、机构同现及作者同现图。如图 1(a)、(b)分别为系统爬取的文章自动提取的摘要展示界面及无人机在疫情应用中的关键词云图模块。

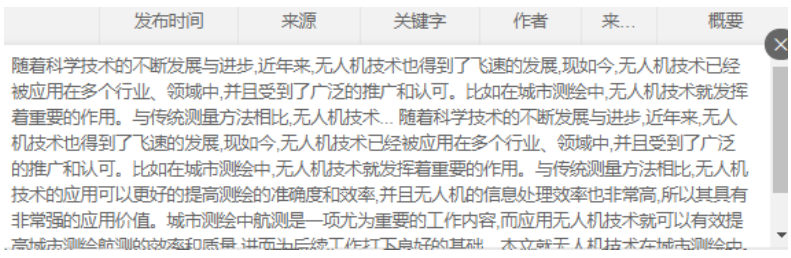



图 1(a) 系统爬取的文章自动提取的摘要展示界面




图 1(b) 无人机应用于疫情中的关键字云图

## 四. 主要研究成果及应用情况

1、**已发表 EI 会议论文 2 篇：**①**Chen Q J,** He Y R, Su Q, et al. Building A Natural Disaster Knowledge Base Expert System based on the Distributed and Incremental Crawling Technology[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2020, 435(1): 012-024.(EI 检索)；②**Chen Q J**, He Y R, He T T, et al. The Typhoon Disaster Analysis Emergency Response System Based on UAV Remote Sensing Technology[J]. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, 42: 959-965.(EI 检索)；2、**在 SCI 期刊《Microelectronics Reliability》已发表一篇论文**：Wu K S, He Y R,**Chen Q J**,et al.Analysis on the damage and recovery of typhoon disaster based on UAV orthograph[J].Microelectronics Reliability,2020,10(7):1-11. (SCI 检索)；

**3、申请发明型专利 1 项：**何原荣,**陈秋瑾**,苏群,等.知识库构建方法、装置和知识搜索方法、系统[P].发明专利,专利号:202010373034.0（已过初审，实审中）；**4、申请软件著作权 1 项：**何原荣,**陈秋瑾,**苏群.自然灾害知识库专家系统(简称:灾情知识库系统)V1.0.登记号:2020SR02200217；5、项目研究成果目前在数字福建自然灾害大数据研究所得到成功应用。

**联系人：陈秋瑾　　电　话：18506927528　　Email:892725709@qq.com**