



武汉大学数据科学与智能应用实验室

Data Science and Intelligent Application Lab

Language
Information
Financial
Large
Management
Knowledge
Graph
Model
Technology

 hong@whu.edu.cn

 武汉大学

目录

CONTENTS

1. 团队介绍

2. 项目简介

3. 科研成果

4. 科研产品

数据科学与智能应用实验室

团队简介

武汉大学数据科学与智能应用实验室（Data Science and Intelligent Application Lab，简称DASI）武汉大学人工智能学院和武汉大学大数据研究院等单位建立，负责人为洪亮教授。目前，实验室研究团队包括博士后、博士生，硕士生，本科生共20余名。

负责人简介

洪亮（1982-），武汉大学人工智能学院教授，博士生导师，现任武汉大学大数据研究院院长助理、技术研发中心主任。2017年入选武汉大学珞珈青年学者，2021年入选武汉大学人文社会科学优秀青年学者。洪亮于2009年毕业于华中科技大学计算机学院，获工学博士学位，弗吉尼亚大学（University of Virginia）计算机系联合培养博士；2012年担任微软亚洲研究院访问研究员；2015-2016年南加州大学（University of Southern California）计算机系访问学者。

在TKDE、TKDD、KDD、SIGSPATIAL、管理世界、管理科学学报等国内外高水平学术期刊和会议上发表论文60余篇，并获得10余项国内外专利授权，撰写中英文专著多部。2016年获得湖北省自然科学优秀学术论文二等奖，2021年获得全国高校教学创新大赛三等奖，2023年获得高等教育国家级教学成果奖一等奖。先后承担国家重点研发计划，国家自然科学基金重点项目、面上项目、青年项目，教育部博士点基金，湖北省重点研发计划等各类国家级、省部级和企业科研项目。担任KDD、ICDE、IJCAI、TKDE，TIST等CCF A类会议和期刊的程序委员会委员和审稿人。现为中国计算机学会(CCF)高级会员，CCF数据库专业委员会执行委员，ACM、IEEE会员。

指导学生获得“花旗杯”金融创新应用大赛全国二等奖，荣获“优秀指导教师”称号。指导学生获得中国研究生金融科技创新大赛全国二等奖，JCDL 2020最佳实践论文奖。指导的多个大学生创新创业训练计划项目获批国家级立项。



研究方向

本实验室聚焦 图基础模型（Graph Foundation Model）与 多模态RAG（Retrieval-Augmented Generation）的交叉研究，致力于解决知识图谱构建与分析中的关键瓶颈问题。通过构建统一图分析架构提升跨领域泛化能力，结合多模态增强检索技术突破数据稀疏与人工依赖限制，形成“结构解析-语义生成”双驱动引擎。研究方向深度融合金融时序分析、跨模态数据融合等前沿技术，为金融风控、数字人文等场景提供可复用的智能化基础设施，推动知识计算从静态构建向动态推理的范式升级。

【图基础模型】在图基础模型方向，实验室突破传统图分析模型的任务特异性限制，构建支持跨任务（节点分类、链路预测、子图识别）和跨领域复用的统一架构。关键技术包括：基于图神经网络的共享结构特征提取，实现多源异构时序金融数据的动态模式挖掘；融合金融/医疗领域本体与规则的领域知识嵌入，增强复杂场景适配性；面向稀疏数据的高效推理优化，支撑实时风险推演与跨市场关联分析。应用层面，已构建亿级金融时序知识图谱平台，在深圳证券、交通银行等机构实现股权穿透、舆情监测、信贷风控等场景验证，并可通过微调技术向医疗网络分析、政务关系挖掘等领域迁移。

【多模态RAG】实验室着重攻克传统信息处理模式在面对复杂、多元数据时所暴露出的单模态局限性与检索精准度不足等难题，打造多模态融合的高效检索 - 生成架构。关键技术体现在：借助先进的跨模态特征提取技术，打破文本、图像、音频等不同模态数据间的壁垒，实现深度语义关联，为精准检索奠定基础；设计智能检索策略，依据任务需求动态调整检索范围与粒度，从海量多模态知识库中快速定位关键信息；构建多模态协同生成机制，使语言生成模型在文本生成过程中能有机融入图像描述、音频情感等多模态元素，提升生成内容的丰富度与贴合度。在应用场景拓展方面，于金融风控里，多模态 RAG 能融合交易文本记录、监控图像等多源数据，更精准识别风险信号；在数字人文领域，可结合古籍文字、文物图像等多模态资料，助力文化研究与知识挖掘，为多领域提供高效、智能的多模态信息处理解决方案，与图基础模型协同发力，赋能知识计算动态升级。

目录

CONTENTS

1. 团队介绍

2. 项目简介

3. 科研成果

4. 科研产品

项目简介

数智时代，许多大型企业具有深厚的知识积累，但是在知识应用层面还存在若干提升优化空间。项目旨在利用数字技术解决企业在知识管理与应用层面面临的问题，特别是数字技术应用滞后、知识流失与传承困难、员工与企业发展亟需数字赋能等问题。通过引入大语言模型和知识图谱技术，项目致力于建立智能化的知识管理系统，提升企业的知识应用效率和员工生产力。

研究内容

□数字技术与知识管理的融合：

- **问题:** 传统知识管理系统依赖大量人力，效率低下；员工知识流失严重，新员工培养成本高；企业和员工亟需数字技术赋能。
- **方案:** 引入“知识图谱+大模型”人工智能技术，构建知识中台，以智能化手段将海量数据转换为结构化知识，形成完整的企业知识图谱。

□联邦型分布式知识图谱管理方法：

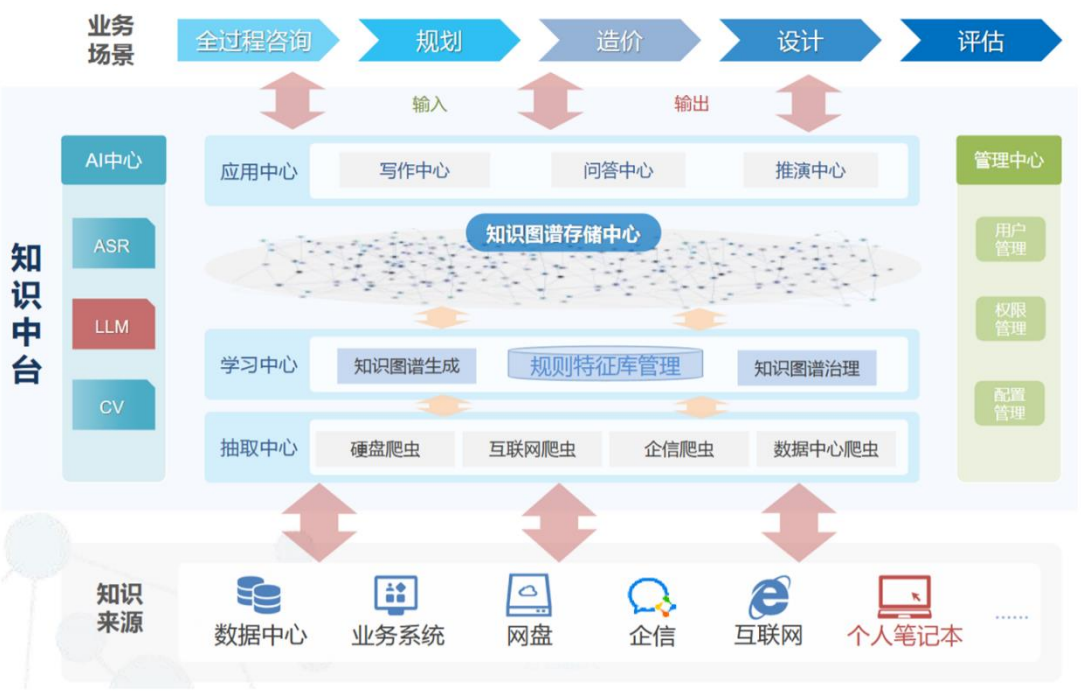
- **目标:** 通过智能手段实现知识的高效管理与应用。
- **方法:** 搭建知识中台，利用大语言模型与知识图谱技术，构建企业知识图谱，实现知识的抽取、融合与应用。

□金融图立方可视化分析技术：

- **功能:** 利用机器人和人机协同，提高设计咨询人员的效率。
- **方案:** 在知识中台与知识图谱的基础上，打造数字化设计咨询智能助手和虚拟顾问团，通过机器人、人机协同全方位、全过程赋能设计咨询人员，形成覆盖公司的知识抽取、知识融合、知识应用的全流程核心能力，推动公司知识全面活化。

新型电力系统知识图谱与大语言模型融合技术研究与应用

--横向项目



项目成果

- 利用智能化手段，将海量数据转化为结构化知识，构建完整的企业知识图谱。
- 通过数字化设计咨询智能助手和虚拟顾问团，实现了知识管理的全流程覆盖。
- 项目实施过程中，将大语言模型与知识图谱技术深度融入日常工作流程。
- 构建知识中台和企业知识图谱，并在此基础上开发数字化设计咨询智能助手和虚拟顾问团。

项目简介

定密工作是保密工作的基础性和源头性工作，事关国家安全与发展。要实现智能化定密需要领域专家的规则知识进行补充，并对知识库进行迭代式更新。本项目旨在利用智能化技术解决保密工作中定密标注的难点和问题。通过构建智能化密点标注辅助系统，利用知识图谱技术提升定密工作的准确性和效率，确保国家安全和保密工作的科学化、精准化。

技术目标

□构建密点知识图谱：

- 将非结构化和半结构化数据进行结构化处理，建立密点知识图谱。
- 知识图谱作为外部知识库，通过信息检索、知识推理和分析技术支持定密工作。

□智能化密点标注功能：

- 利用大型预训练模型进行密点标注，考虑上下文语义和词语的同义词识别。
- 提升定密准确性，识别人工定密难以发现的风险点，支持决策者的科学决策依据。

□知识图谱的推理能力：

- 提供对标注密点的可解释性，解释定密标注的依据和逻辑。
- 支持定密经验的积累和知识库的迭代更新，适应保密管理领域的动态变化和复杂性。

智能化密点标注辅助系统研究及测试项目

--横向项目



项目成果

- 结合多源数据和时间图技术，实现了对涉密人员行为的全面监控和异常检测。
- 通过图查询和图模式匹配技术，识别复杂的异常行为模式，提高了风险预警的准确性。
- 基于专家赋权和减分模式的评分机制，提供了涉密人员风险评估的科学依据和决策支持。
- 建立了基于时间图的SM人员异常风险预警算法，有效识别涉密人员的异常行为和风险特征。
- 提升了涉密人员管理的科学化和精准化水平，为国家安全提供了重要技术保障。

项目简介

间谍活动是一种为境外敌对势力窃取情报的行为，具有隐秘性高、渗透力强等特点，会对国家安全造成极大危害。因而，如何利用涉密人员信息、行为特征结合关键技术对其风险进行评估是本研究的主要内容。本项目旨在通过建立一套基于时间图的风险评估特征指标，对涉密人员的异常行为进行有效预警和风险评分，从而降低窃密、泄密事件的发生概率。

研究内容

□多源数据利用与时序图构建：

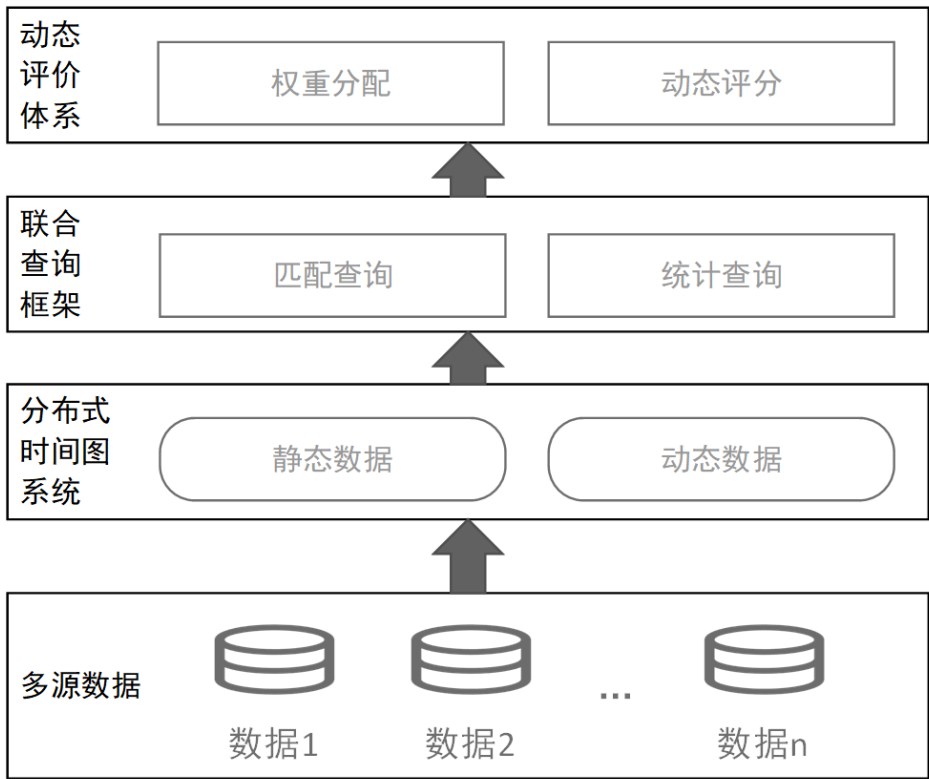
- 利用单位内部各个监管系统中的分散数据，在不影响业务管理的情况下进行联邦分布式管理。
- 构建基于多源数据的时序图，为图分析和异常检测提供数据支撑。

□保密系统自监管：

- 识别涉密人员的异常属性（如高频保存操作、高量保存、异常登录等），利用图查询技术进行行为属性识别。
- 挖掘设计针对涉密人员越权、超范围知悉等异常行为的图模式，利用图模式匹配进行识别。

□涉密人员评分机制：

- 结合涉密人员的基本信息（不可变动信息）和行为信息（可变动信息），进行不同评分模式的评估。
- 通过专家赋权确定指标权重，采用减分模式，在基本信息所产生的基本分数基础上，根据不同异常行为削减相应分数，形成综合风险评分。



项目成果

- 结合多源数据和时间图技术，实现了对涉密人员行为的全面监控和异常检测。
- 通过图查询和图模式匹配技术，识别复杂的异常行为模式，提高了风险预警的准确性。
- 基于专家赋权和减分模式的评分机制，提供了涉密人员风险评估的科学依据和决策支持。
- 建立了基于时间图的SM人员异常风险预警算法，有效识别涉密人员的异常行为和风险特征。
- 提升了涉密人员管理的科学化和精准化水平，为国家安全提供了重要技术保障。

基于科学共同体知识大图的隐性合作关系发现与深度挖掘

--国家自然科学基金面上项目

项目简介

科研合作的多样性使得科学共同体中广泛存在无法通过数据直接反映的隐性合作关系。从知识关联视角，隐性合作关系是学者之间通过知识的流动和转化形成的多类型、多层次知识关联。针对当前研究体系化不足，缺少大数据技术的问题，本项目基于知识关联理论，从“5W1H”六个方面深入研究科学共同体隐性合作关系发现和深度挖掘的技术。

首先对隐性合作关系进行规范化的定义和分类，研究隐性合作知识组织和表示模型，以及人机协同迭代式的知识大图构建方法，基于学术大数据构建科学共同体知识大图；

然后研究基于知识推理的隐性合作关系发现方法，通过知识大图中的知识关联发现隐性合作关系的知识关联；

在此基础上基于社区画像揭示隐性合作方式并进行对比分析，基于特征抽取探测隐性合作动机并建立影响模型，最后基于时空模式分析和建模隐性合作的时空演化。

研究进展

□科学共同体知识大图概念模型：

利用多重语义蕴涵图对多层次、多角度的知识关联进行准确地组织和表示，基于知识大图拓展知识关联的角度与层次的表示，实现了知识的全局关联。

□基于提示集成的隐性合作关系发现：

设计融入关系语义的提示模板，增强模型提示效果；通过多模板集成进行数据增强，缓解少样本问题；引入基于原型网络的标签映射方法，并通过实例级注意力机制提升分类准确性。通过融合多源异构学术大数据构建科学共同体知识大图，共包括2,092,356篇论文，8,024,869条论文之间的引用关系，1,712,433位作者信息，以及4,258,615条合作关系。

基于科学共同体知识大图的隐性合作关系发现与深度挖掘

--国家自然科学基金面上项目

研究进展

□基于规则标注的隐性合作动机探测：

抽取和建模规则语义关联，设计标签感知的规则生成方法，采用交叉注意力机制进行语义匹配，并通过不一致性导向的主动学习来验证规则。此方法有效缓解了语义漂移问题，提升了隐性合作动机探测的准确性。

□科学共同体知识大图隐性合作关系推理：

在选择最佳搜索节点时，加入优先级以解决不同关联关系的重要性差异问题，在计算路径权重时，引入启发式评估函数以减少路径搜索的时间与空间代价。

项目成果

- 洪亮, 马费成. 面向大数据管理决策的知识关联分析与知识大图构建[J]. 管理世界, 2022, 38(1): 207-219.
- Hou W, Hong L, Xu H, et al. RoRED: Bootstrapping labeling rule discovery for robust relation extraction[J]. Information Sciences, 2023, 629: 62-76. (SCI—区)。
- Hou W, Hong L, Zhu Z. KGRED: Knowledge-graph-based rule discovery for weakly supervised data labeling[J]. Information Processing & Management, 2024, 61(5): 103816. (SCI—区)。
- 徐豪帅, 洪亮, 侯雯君. 基于提示集成的少样本关系抽取方法[J]. 数据分析与知识发现, 2024: 1.
- 洪亮(1/6); 武汉市社会科学优秀成果奖, 武汉市人民政府, 中共武汉市委, 2022.
- 张震; 缪亚男; 马欢; 陆伟; 洪亮; 邓诗琦; 一种面向智能应用的领域本体构建方法, 2022-8-30, 中国, ZL201811307870.8 (专利)
- 武汉大学; 面向信息抽取的人机协同数据标注平台 [简称: CrowdIE] V1.0, 2023SR0512951, 原始取得, 全部权利, 2023-2-13 (软件著作权)

金融时序知识图谱查询与分析平台及应用验证

--科技创新2030-“新一代人工智能”重大项目

项目简介

不同领域的金融机构不断产生海量数据，数据的种类多样、格式复杂，随时间不断演化，如何对这些数据采集和汇聚是当前面临的难点问题。本项目旨在通过研究金融跨领域数据汇聚技术、联邦型分布式知识图谱管理方法及金融图立方可视化分析技术，建立一个高效的金融时序知识图谱查询与分析平台，构建金融风控大脑，支持跨领域金融风险识别、监测、预警与防控。

研究内容

□金融跨领域数据汇聚技术：

- 解决多源异构金融数据的采集和汇聚难题。
- 通过多系统集成技术和分布式框架，实现对大规模离线数据和实时数据的高效接入。
- 提供统一访问接口，保障数据的全程管理及安全接入。

□联邦型分布式知识图谱管理方法：

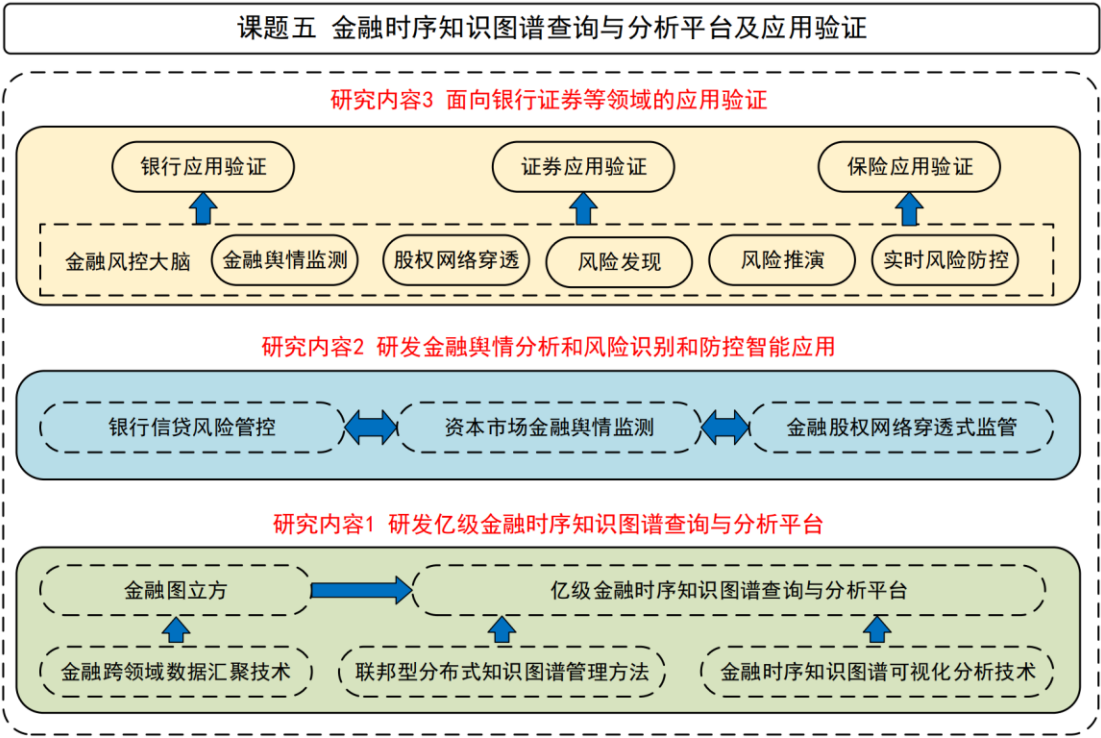
- 研究高效的联邦型分布式存储与管理，支持跨领域数据汇聚和知识图谱的联邦型存储。
- 确保大规模、高并发、强实时的分布式查询与分析能力。

□金融图立方可视化分析技术：

- 针对图数据的弱模式特点，研究时序超图的结构和语义特点。
- 实现交互式、变粒度的金融图立方摘要与可视化分析，提升其可用性。

金融时序知识图谱查询与分析平台及应用验证

--科技创新2030-“新一代人工智能” 重大项目



项目成果

- 通过规范化、体系化的知识分类，解决了跨领域金融数据的汇聚难题。
- 联邦型分布式存储与管理方法确保了数据的高效存储和实时查询分析。
- 金融图立方可视化技术大幅提升了金融数据的可用性和分析效率。
- 项目开发的智能应用包括银行信贷风险管控、资本市场舆情监测和金融股权穿透式监管等。
- 已在5家金融机构（深圳证券信息有限公司、交通银行、中国建设银行、恒丰银行、众邦银行）进行了典型应用验证，涵盖了上述智能应用场景

基于知识大图的中小企业信贷风控

--横向项目

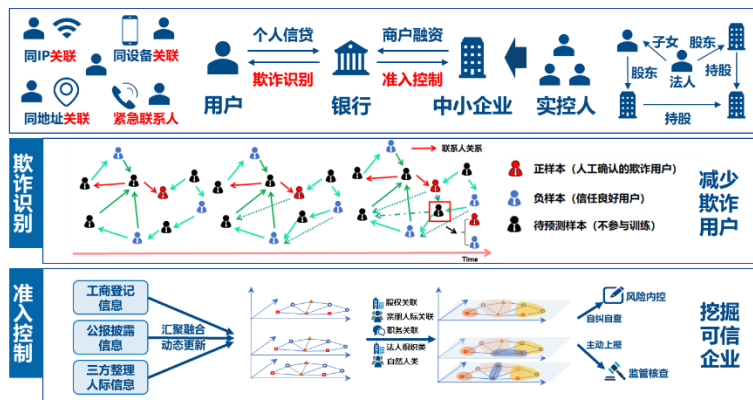
项目简介

本项目旨在利用C端知识图谱和图卷积神经网络（GCN）模型，提升零售风控数据处理和信贷欺诈检测的效率和准确性。通过构建高效的知识图谱，优化数据查询、清洗和分析过程，从而增强反欺诈模型的表现，并有效扩展高信用企业白名单。

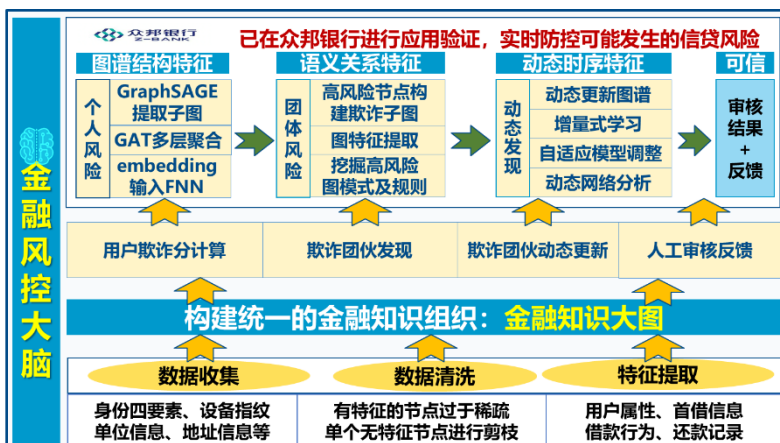
研究内容

数据结构与图谱构建：

- 底表数据结构：**以订单号为主键，存储客户信息、账户信息、联系人信息、工作单位信息、设备指纹、申请时IP与GPS等。
- 知识图谱构建：**建立实体与图谱结构模型，实现历史数据和实时数据的自动对齐，提升关联查询效率，简化清洗规则和衍生指标计算。



两个业务均基于人工核查或者简单规则，利用知识大图实现降本增效



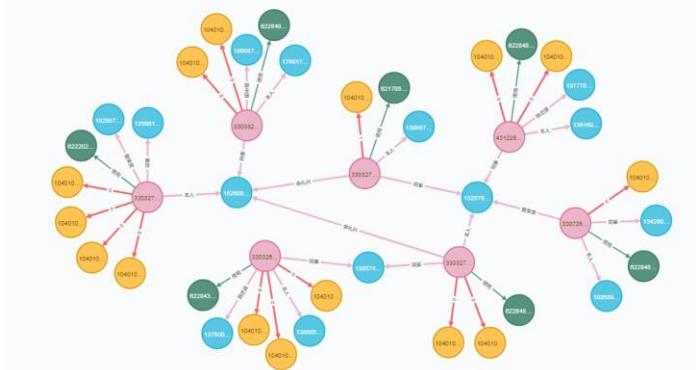
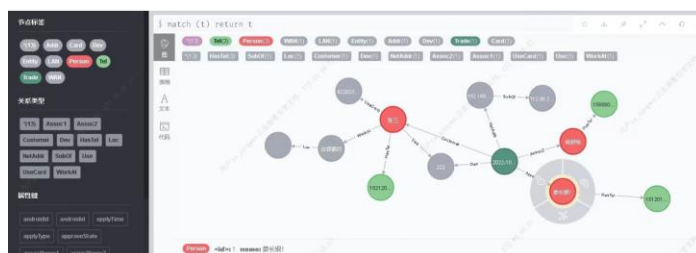
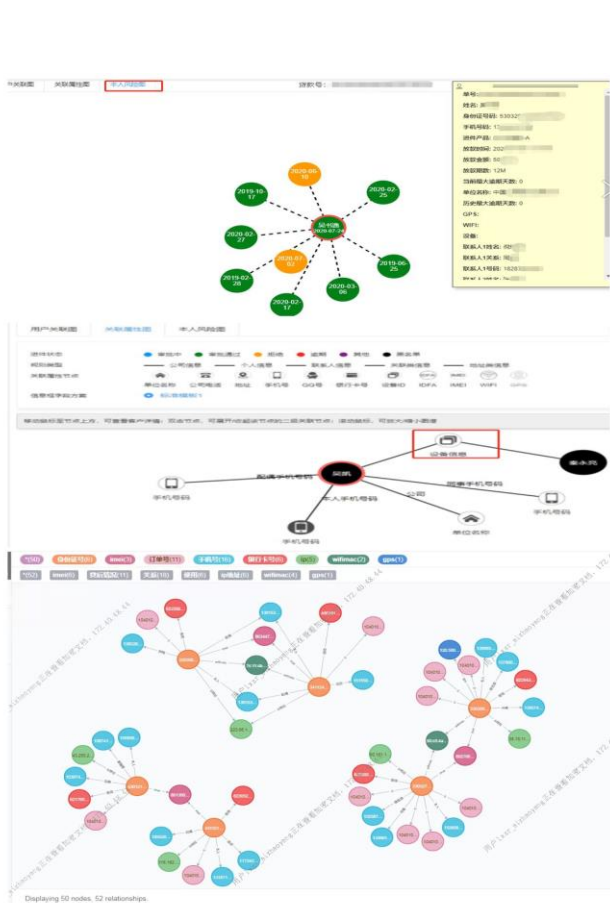
基于知识大图的中小企业信贷风控

--横向项目

研究内容

GCN模型应用:

- 特征工程**: 处理节点的缺失值、增加节点度数、使用时间戳信息和相似度信息; 为边增加反向边、构造边方向性信息及编码。
- 模型训练**: 使用降采样平衡样本数量, 利用集成方法提高模型稳定性。
- 数据处理**: 在1505574个节点和684512条有向边的图上进行训练和预测, 仅在成功申请的节点上进行分类或回归, 其余节点作为特征信息使用。



操作/输入	输入/输出	模型名称	模型输出
score	→	普普复神经网络模型V1	S_ZB_VB_ZYFZWLMODELV1
F_ZB_VB_APPLY3SCORE	→	联系人2图评分_武大联合版	F_ZB_VB_APPLY3SCORE
F_ZB_VB_APPLYSORE	→	申请人图评分_武大联合版	F_ZB_VB_APPLYSORE
F_ZB_VB_APPLY2SCORE	→	联系人1图评分_武大联合版	F_ZB_VB_APPLY2SCORE

项目成果

- 高效的知识图谱构建显著提升了数据查询和处理效率。
- GCN模型稳定且高效，优化了信贷欺诈检测过程。信贷欺诈检测模型效果良好：AUC=0.6295，KS=0.1096；在Dgraph公开数据集上，AUC=0.8455，KS=0.5543。模型充分利用图结构和时序信息，无需繁琐特征工程。
- 大规模白名单扩展，有效支持业务发展和风险控制。由1455家核心企业扩展至49396家白名单企业，扩展比例达34倍，支持核心客户开发，减少信贷风险。

项目简介

随着电信诈骗监管趋严，不法分子不断研究新型作案方式，利用出租出借“跑分”掌握一批黑灰产账户，黑产链条趋于产业化、隐蔽化。本项目致力于解决随着电信诈骗复杂化和产业化趋势，传统手段难以应对的问题。通过建立多源异质数据的联合分析模型，利用知识大图技术深入挖掘和分析电信诈骗背后的关联网络和趋势变化，以提高涉赌涉诈账户的识别准确度和覆盖率，支持风控运营的有效实施。

研究内容

□知识大图构建：

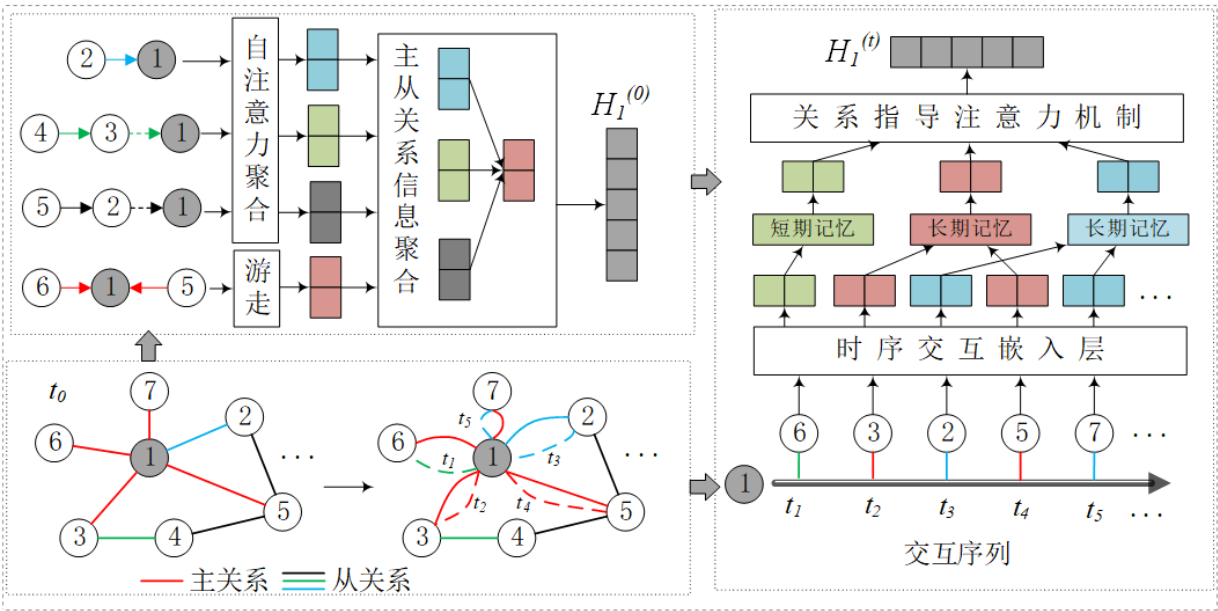
- 基于建信金科BigData Center的知识图谱建模方法与工具，构建以小微企业、客户等节点为基础的点边关系底图。
- 扩展黑灰产账户特征维度，利用算法建模，构建完整的知识大图。提供统一访问接口，保障数据的全程管理及安全接入。

□涉诈类知识图谱账户挖掘模型：

- 描绘欺诈团伙的关联网络，利用风险评分提高涉赌涉诈账户的识别准确度。
- 输出诈骗团伙名单数据，为行方反赌反诈模型提供迭代依据，提升风控运营中涉赌涉诈账户的管控率。

技术方案

- 采用表示学习模型，从资金流转关系和其他关系中捕获不合理资金流转信息。
- 基于长短记忆模型建模和异质交互聚合的时序知识图谱表示学习模型，实现黑灰账户的时序识别与分析。

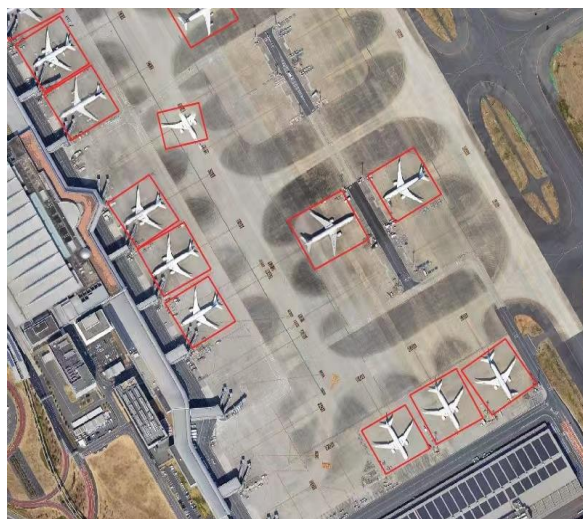


项目成果

- 结合大数据与人工智能技术，构建了高度智能化的黑灰产识别系统。
- 利用时序知识图谱表示学习模型，有效应对电信诈骗作案方式的时变性和复杂性。
- 在实际应用中取得了显著的成效，为金融行业的风险控制提供了前沿技术和创新解决方案。

技术方案

- 采用深度学习和迁移学习技术，实现目标的语义信息提取。
- 采用Transformer技术，实现地物要素的提取，感知环境要素目标。



项目成果

- 结合大数据与人工智能技术，构建目标语义感知识别系统。
- 结合大模型技术，构建地物要素语义分割模型。

目录

CONTENTS

1. 团队介绍

2. 项目简介

3. 科研成果

4. 科研产品

Disclosing Actual Controller based on Equity Knowledge Graph Learning

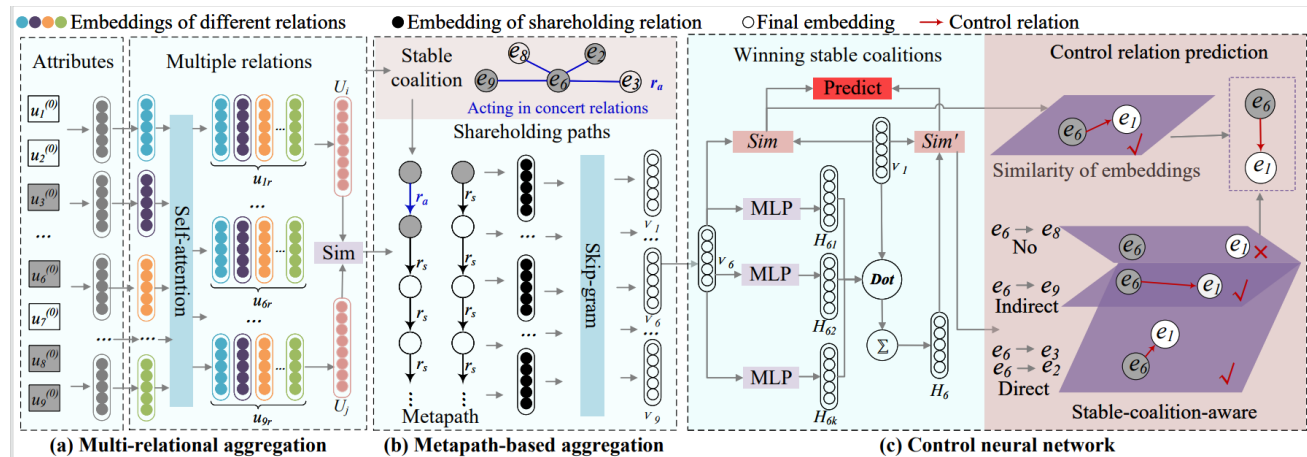
Qingying Xu, Liang Hong*, MingXuan Shen, BaoKun Yi

期刊信息

- 2024年第31届国际知识发现与数据挖掘大会（ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD）的数据科学（Applied Data Science, ADS）分会收录此论文。
- KDD是中国计算机学会（CCF）推荐的A类国际学术会议，为数据挖掘领域的顶级会议。本次ADS分会投稿量达285篇，仅22%被录用。

研究内容

- 构建一个股权知识图谱来表示股权网络的复杂语义和结构信息，并将实际控制人挖掘任务转化为股权知识图谱上控制关系预测。
- 为识别获胜稳定联盟，提出了一种基于股权知识图学习方法（EKGL）。
- 为发现股东间稳定的联盟关系，在EKGL中设计了一个多关系聚合模块，通过综合股东间不同类别的关系，预测股东一致行动概率。
- 利用一种元路径聚合模块，捕获股东的持股路径，对持股结构进行编码，从而衡量股东的持股比例。
- 为确定获胜的稳定联盟，设计了控制神经网络来模拟股东的投票过程。
- 基于真实数据集构建的股权知识图谱的实验和案例研究验证了EKGL的有效性。



Discovery and Analysis of Control Associations in Shareholding Networks Based on the Voting Game Approach

Zhonglu Shi, Liang Hong*, Chenyu Zhao

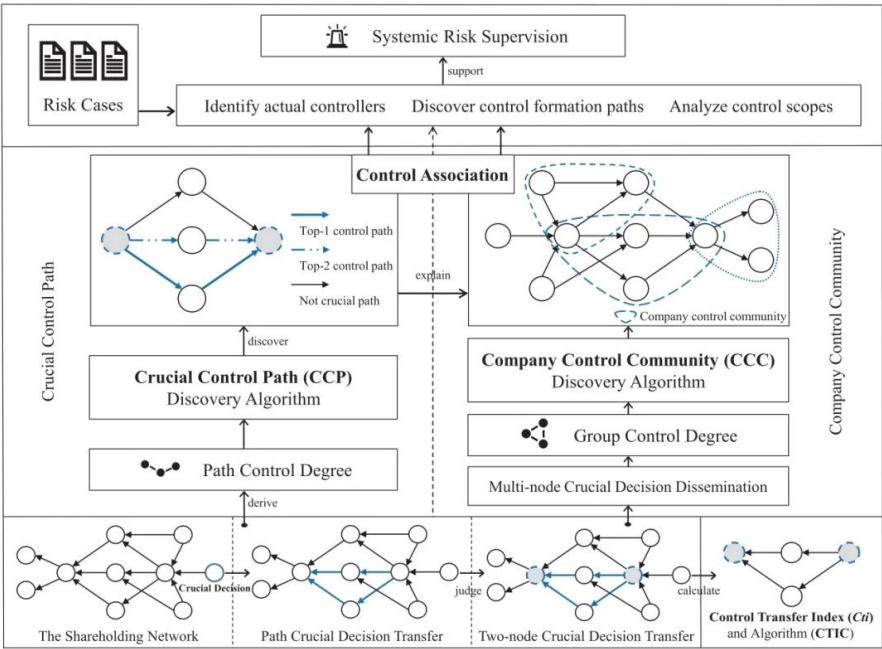
本研究工作得到科技创新2030-新一代人工智能重大项目（批准号：2020AAA0108505）、国家自然科学基金重大项目（批准号：7207417）的资助。

期刊信息

- 2023年，IEEE ACCESS期刊录用此论文。
- IEEE ACCESS 为JCR Q2期刊，影响因子(2021)为3.476，5年影响因子为3.758。

研究内容

- 从投票博弈角度考虑包括最终持有人以及中间结点在内的控制作用，提出控制传递指数（Cti）以衡量所有公司之间控制关系强弱程度。
- 提出控制权转移指标计算（CTIC）算法来精确计算Cti。
- 提出关键控制路径（CCP）概念及发现算法，发现公司之间更重要的控制链路。
- 提出公司控制社区（CCC）概念及发现算法，揭示满足控制度阈值的重叠公司社区。
- 针对以上关键控制关联进行控制与风险案例分析，说明这些概念及算法在识别公司实际控制人、发现控制形成路径和分析控制范围方面的有效性。



(股权网络控制关联发现与分析框架)

RoRED: Bootstrapping Labeling Rule Discovery for Robust Relation Extraction

Wenjun Hou, Liang Hong*, HaoShuai Xu, Wei Yin

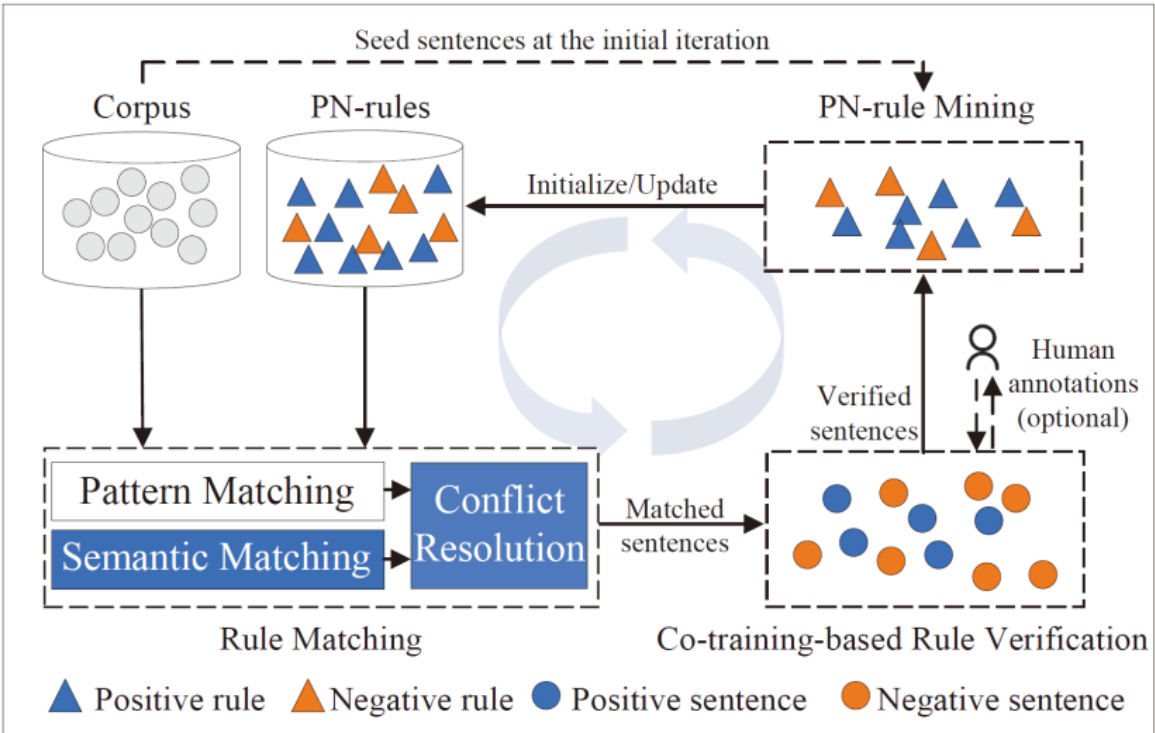
本研究工作得到国家自然科学基金“基于科学共同体知识大图的隐性合作关系发现与深度挖掘”(基金号72074172)的资助。

期刊信息

- 2023年, Information Sciences期刊录用此论文。
- Information Sciences是信息系统领域权威期刊, 中科院、JCR一区TOP期刊, 2022年最新影响因子为8.233。

研究内容

- 提出一种自举式规则发现方法以增强关系抽取的鲁棒性, 来降低弱监督关系抽取任务中由错误标签和遗漏标签带来的噪音。
- 设计PN规则, 通过正负规则协同过滤错误标签, 提升规则匹配的精确率。
- 提出基于词-句语义关联网络的语义匹配方法, 通过词-句子-规则之间的语义关联来发掘遗漏标签, 提升规则匹配的覆盖范围。
- 提出基于协同训练的规则验证机制, 对句子的标签进行验证, 以提高规则发现的整体质量。
- 基于实验结果验证了该工作的有效性。



(自举式规则发现框架)

KGRED: Knowledge-graph-based rule discovery for weakly supervised data labeling

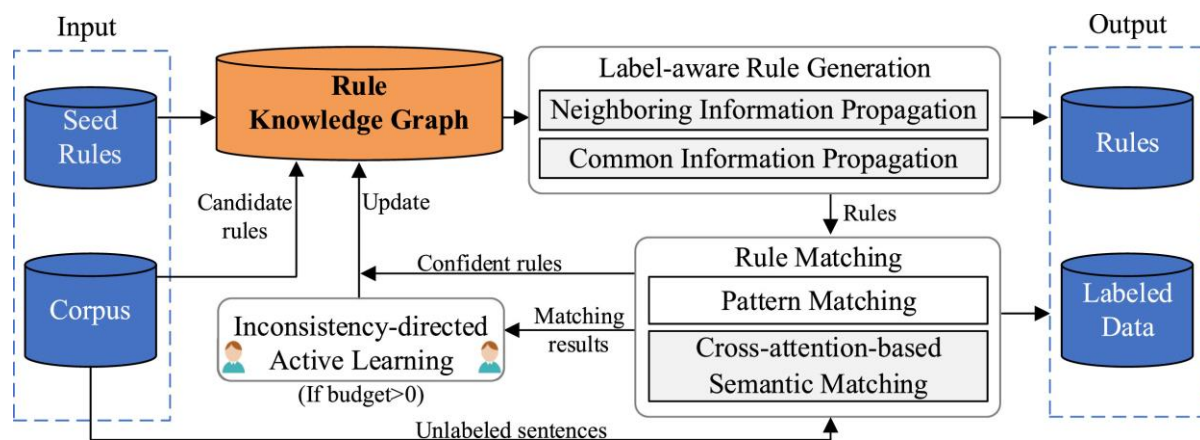
Wenjun Hou, Liang Hong*, Ziyi Zhu

期刊信息

- 2024年, Information Processing & Management期刊录用此论文。
- Information Processing & Management是信息系统领域权威期刊, 中科院、JCR一区TOP期刊, 2023年最新影响因子为7.4。

研究内容

- 提出一种基于知识图的规则发现方法 (KGRED), 利用规则之间的多维语义关联来减轻规则发现中的语义漂移。
- 为减少错误分类的规则, 设计了一种标签感知规则生成方法, 专注地将先验知识从种子规则传播到基于规则 KG 的候选规则。
- 为减少错误匹配的句子, 提出了一种基于交叉注意力的语义匹配机制, 以细化句子的语义信息, 同时丰富规则的语义信息。
- 提出一种不一致导向的主动学习策略来验证在规则生成和匹配中表现不一致的规则。在两个公共数据集上的实验证明, 与最先进的方法相比, KGRED可以实现至少5.1%的F1分数提高。



(KGRED框架)

面向大数据管理决策的知识关联分析与知识大图构建

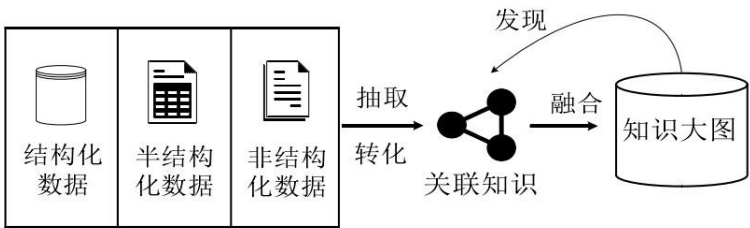
洪亮，马费成

期刊信息

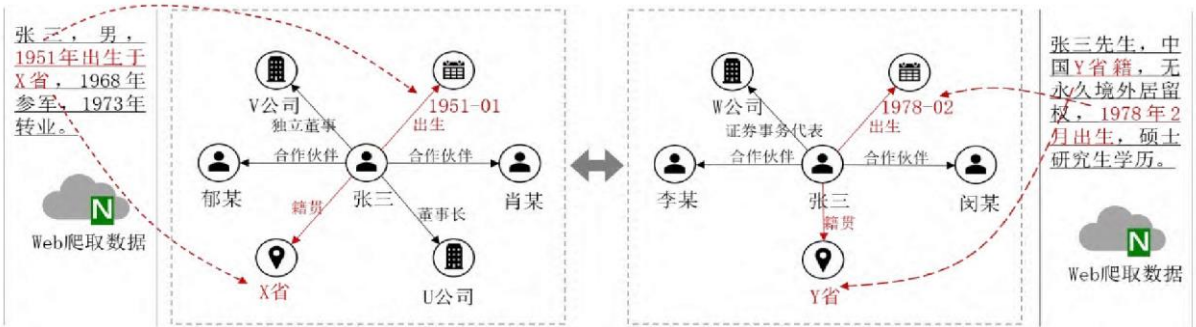
- 2022年，管理世界期刊录用此论文。
- 《管理世界》(Management World)是由国务院发展研究中心主管、国务院发展研究中心主办的中文版月刊。该刊被CSSCI、北大核心合集等数据库收录，(2024版)复合影响因子为35.785，(2024版)综合影响因子为27.332。

研究内容

- 对知识关联分析和知识大图构建问题进行了形式化的定义，利用多重语义蕴涵关系表示大数据中存在的多角度、多层次的知识关联。
- 提出面向大数据管理决策的知识关联分析与知识大图构建统一框架和体系化方法。
- 通过典型应用场景：股权知识大图穿透式监管与疫情知识大图精准溯源，对所提出方法的有效性进行了验证。



(知识大图构建框架)



(关联知识融合示例)

基于提示集成的少样本关系抽取方法（录用定稿）

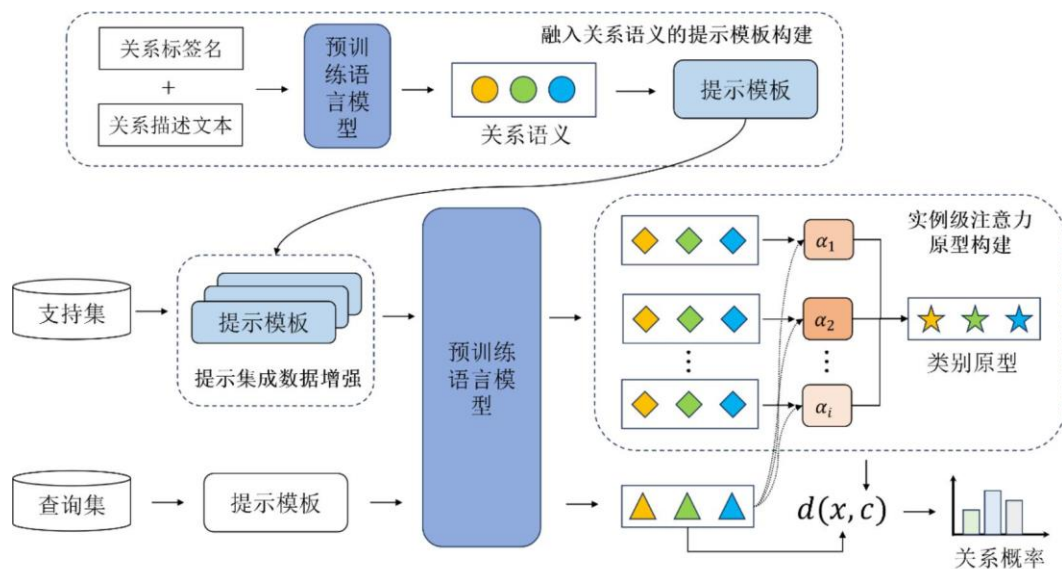
徐豪帅, 洪亮*, 侯雯君

期刊信息

- 2023年，数据分析与知识发现期刊录用此论文。
- 《数据分析与知识发现》（Data Analysis and Knowledge Discovery）是由中国科学院主管、中国科学院文献情报中心主办的学术性专业期刊。期刊已被国内各主要核心期刊目录及引文来源库收录，入选《FMS管理科学高质量期刊推荐列表》，并被美国工程索引数据库（EI Compendex）和SCOPUS收录。（2024版）复合影响因子为3.957，（2024版）综合影响因子为2.385。

研究内容

- 提出一种基于提示集成的少样本关系抽取方法。该方法利用训练样本的回答词表示构建关系类别原型，通过计算关系类别原型与回答词表示之间的距离来构建标签映射关系。
- 提出一种关系语义注入的提示模板构建策略，通过预训练语言模型对关系标签文本和关系描述文本的语义进行编码，将其融入到提示模板中，增强模板对预训练语言模型的提示效果，使得在原型的构建过程中融入关系标签的语义信息，丰富回答词原型所包含的信息。
- 提出一种提示集成数据增强的方法，对单个关系上下文实例应用多个提示模板，以提示集成的方式对数据进行增强，缓解了少样本场景下原型构建所包含的样本信息有限的问题。
- 引入了一种实例级注意力机制，在原型构建的过程融入与任务更相关的特征，降低异常样本对原型构建的影响，增强原型构建过程鲁棒性，提升原型表示的准确性。



(方法框架图)

知识关联视角下的文化遗产知识大图研究：理论、方法和趋势

洪亮，宋睿，朱丽雅，侯雯君

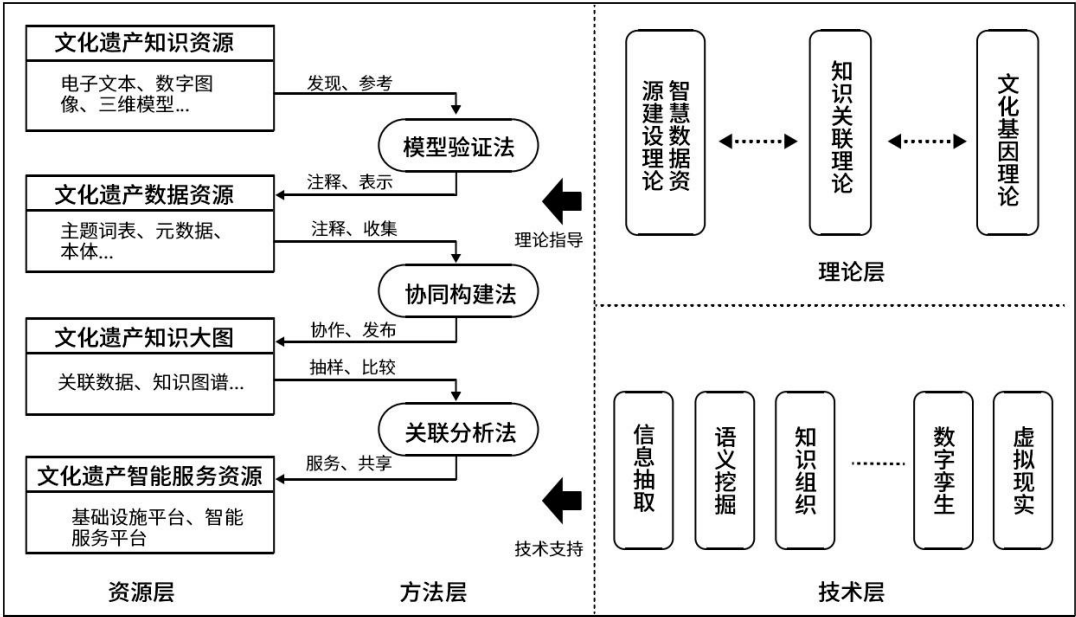
本研究得到国家自然科学基金面上项目“基于科学共同体知识大图的隐性合作关系发现与深度挖掘”（72074172）和湖北省重点研发计划项目“知识图谱与机器学习融合及高级推理关键技术研究”（2020BAB026）的资助。

期刊信息

- 2022年，图书情报知识期刊录用此论文。
- 图书情报知识杂志是北大核心、CSSCI、SCD目录期刊，是一本由武汉大学主办的优秀新闻类期刊。(2024版)复合影响因子为4.37，(2024版)综合影响因子为3.584。

研究内容

- 从知识关联视角，系统梳理和分析不同学科中现有的理论基础，明晰不同理论之间的逻辑关系，从而给出文化遗产知识大图研究的理论框架。
- 对文化遗产知识大图的相关研究方法进行总结，并依据知识大图构建过程中的资源转化流程，建立文化遗产知识大图的研究方法体系。
- 在理论框架和方法体系的基础上，总结未来研究趋势，以期为文化遗产知识大图后续研究提供进一步的参考与借鉴。



(文化遗产知识大图研究的方法体系)

唐诗知识图谱的构建及其智能知识服务设计

周莉娜，洪亮*，高子阳

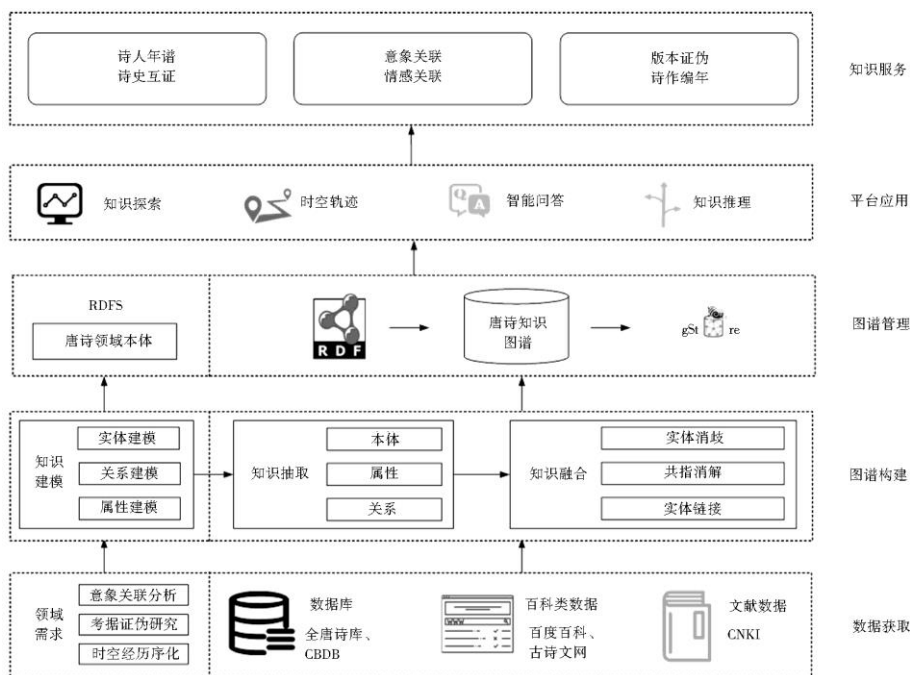
本研究得到国家重点研发计划“科学大数据管理系统”子课题“图数据管理关键技术及系统”(项目编号:2016YFB1000603)和教育部人文社会科学重点研究基地重大项目“大数据资源的语义表示与组织研究——面向文化遗产领域”(项目编号:16JJD870002)的资助

期刊信息

- 2019年，图书情报工作期刊录用此论文。
- 《图书情报工作》是由中国科学院主管、中国科学院文献情报中心主办的国家级大型图书馆学情报学两栖专业学术核心期刊，主要报道以图书馆学、情报学为核心的相关领域理论和实践的最新进展。(2024版)复合影响因子为4.0，(2024版)综合影响因子为3.185。

研究内容

- 设计基于唐诗知识图谱的智能知识服务平台KnowPoetry, 提供唐诗领域的知识探索、时空轨迹、语义查询等智能化知识服务, 推动人工智能环境下唐诗数字人文研究方法创新转型。
- 在对领域知识服务需求调研的基础上, 设计领域知识服务驱动的唐诗本体模型。
- 利用从Web上爬取的多源异构数据, 采用知识抽取、知识融合、知识推理等技术自动构建唐诗知识图谱, 统一表示和组织唐诗领域数据, 实现对大规模唐诗数据的语义化处理。



(唐诗知识图谱构建框架)

金融股权知识大图的知识关联发现与风险分析
洪亮， 欧阳晓凤

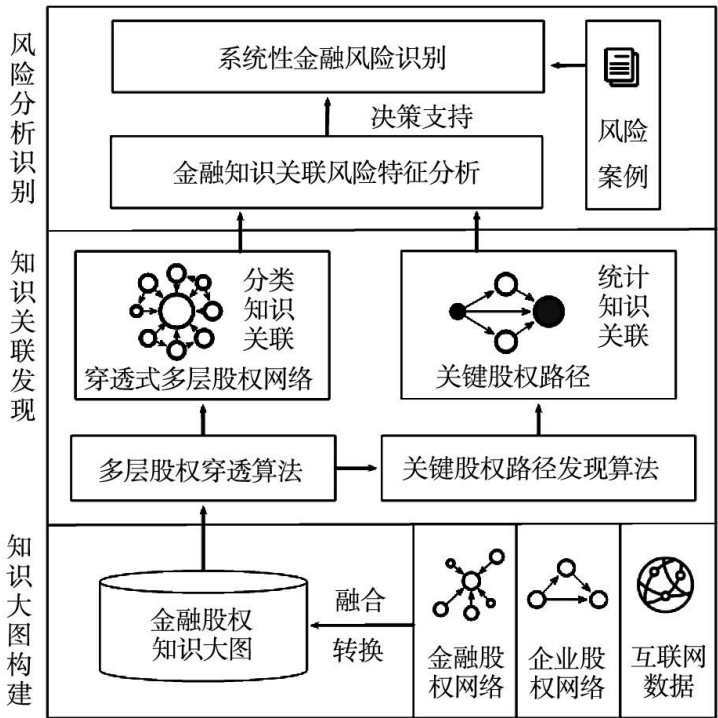
本研究工作得到国家自然科学基金面上项目“基于科学共同体知识大图的隐性合作关系发现与深度挖掘”（72074172）和湖北省重点研发计划项目“知识图谱与机器学习融合及高级推理关键技术研究（”2020BAB026）的资助

期刊信息

- 2022年，管理科学学报期刊录用此论文。
- 《管理科学学报》由中华人民共和国教育部主管，国家自然科学基金委员会管理科学部与天津大学共同主办的管理科学领域类型的学术刊物。(2024版)复合影响因子为5.428，(2024版)综合影响因子为3.734。

研究内容

- 从知识关联理论出发，基于大规模股权网络构建金融股权知识大图。
- 提出多层股权穿透(MEP)算法，揭示以金融机构为中心的穿透式多层股权网络。
- 提出关键股权路径发现(CEP)算法，发现金融机构之间的关键股权路径。
- 结合案例对以上基于知识关联的股权风险结构进行了实证分析，并对MEP和CEP算法进行了性能分析与对比。



(研究框架)

专利申请

软件著作权



发明专利



实验室学子在第十八届“花旗杯”金融创新应用大赛中斩获佳绩

赛事介绍

第十八届“花旗杯”金融创新应用大赛由花旗银行主办，上海市经济和信息化委员会指导，上海市大学生科技创业基金会承办。在政府、教育界、产业界的支持下，“花旗杯”金融创新应用大赛已成为中国金融信息化领域的标杆性赛事。竞赛团队可自由选择自选赛道和“花旗挑战”项目赛道。2023年大赛共有来自国内外222所高校、167支队伍参加。

获奖情况

2023年6月10日，第十八届“花旗杯”金融创新应用大赛决赛在上海举行。我实验室“珞珈山舆情之眼”团队以总评第六名的成绩荣获全国二等奖及指导老师奖。团队指导老师为洪亮教授，骨干成员包括了我实验室保密管理专业2021级硕士生徐青影、2022级硕士生鞠晓慧。

获奖项目

“珞珈山舆情之眼”项目利用知识大图融合关联跨领域金融数据，考虑舆情事件的时序多元语义关联和风险时序传染与衰减，建立基于金融舆情的企业风险画像系统，为商业银行信贷风控提供技术解决方案。项目团队从市场需求、关键技术、应用验证、商业计划和社会价值等方面全面地介绍了参赛项目，并圆满地回答了评委提出的问题，赢得了现场专业评委的好评。



DASI团队荣获“数字金融探索创新优秀案例”

赛事介绍

“金融风险管理”专题论坛由武汉大学信息管理学院洪亮教授担任论坛主席。论坛聚焦于数字金融时代的风险识别、防控与监管，邀请多位金融领域的专家学者、技术创新者及业界领袖，共同探讨金融风险管理方法的转型与升级。洪亮教授表示，通过技术赋能金融风控，能够有效推动行业的智能化和数字化发展，为构建更加透明、高效的金融生态贡献力量。

获奖情况

武汉大学信息管理学院硕士生沈明轩在大会中提交的案例——“基于知识大图的金融风控大脑”，凭借其在金融风险识别与防控方面的创新成果，荣获“数字金融探索创新优秀案例”。

获奖项目

该案例提出了一种基于知识大图的智能化金融风险管理方案，融合了金融知识图谱、股权穿透分析及舆情监测等前沿技术。案例已成功应用于深证信、交通银行、中国建设银行、众邦银行等多个金融场景，显著提升了风险识别精度与效率，帮助金融机构规避了亿元级别的潜在经济损失。大会评审专家认为，该方案不仅在技术上表现突出，更在实践中展示了优越的经济与社会效益。



DASI团队荣获Graph RAG赛题三等奖

赛事介绍

金融风险管理”专题论坛由武汉大学信息管理学院洪亮教授担任论坛主席。论坛聚焦于数字金融时代的风险识别、防控与监管，邀请多位金融领域的专家学者、技术创新者及业界领袖，共同探讨金融风险管理方法的转型与升级。洪亮教授表示，通过技术赋能金融风控，能够有效推动行业的智能化和数字化发展，为构建更加透明、高效的金融生态贡献力量。

获奖情况

在大会同期举办的CCF中国数字金融大会 Graph RAG 竞赛中，武汉大学博士生蒋翱远组织的DASI团队，荣获大赛三等奖。团队成员包括硕士生沈明轩、博士生徐青影、硕士生贾冉以及本科生宋知恒。

获奖项目

DASI团队聚焦于金融领域知识图谱的构建与检索增强生成（Graph RAG）技术的应用，以金融研报为核心数据源，通过设计高效的文档解析和知识图谱生成策略，实现了大模型在金融知识检索和生成能力上的突破。团队的创新成果为金融科技的数字化发展提供了重要参考，并获得了评审委员会的高度评价。



科研奖励荣誉



HongL (1/1) : 最佳论文奖 (GDMA Workshop Best Paper Award) ; 数据库系统高级应用国际会议(DASFAA), 2020
洪亮 (11/12) : 图书情报与档案管理“全球胜任力”人才培养体系探索与实践等奖;教育部, 2022



HongL (3/4; 通讯作者) : 最佳实践论文奖 (The Best Practitioner Paper Award) 国际数字图书馆联合会议 (JCDL) , 2020



洪亮 (1/1) : 第十八届“花旗杯”金融创新应用大赛二等奖指导老师, 2023



洪亮 (1/6) : 武汉市社会科学优秀成果奖, 2022



洪亮 (1/1) : 全国高校教师教学创新大赛三等奖, 2021

目录

CONTENTS

1. 团队介绍
2. 项目简介
3. 科研成果
4. 科研产品

知意-人在环路知识图谱构建平台

平台介绍

知意，针对数据的多源异构的特点，研发知识图谱自动化构建方法和平台，基于人在回路机制，融合规则和大模型实现高效的知识抽取，提高知识图谱构建的质量，实现端到端的中文知识图谱构建，支撑基于知识图谱的智能服务。该系统已在 2023 年 CNCC 中国计算机大会上进行演示。

平台特点

- 端到端。支持从原始文本、本体构建到抽取三元组，导入图数据库进行查询分析的端到端知识图谱全流程构建。3. 主动学习。对于模型难以判断的数据，可通过主动学习来迭代调整模型，提高人机交互的效率，该系统已在 2023 年 CNCC 中国计算机大会上进行演示
- 弱监督。通过结合原型网络的大模型提示学习和自举式标注规则发现来降低数据标注的噪音，提高信息抽取的鲁棒性。
- 主动学习。对于模型难以判断的数据，可通过主动学习来迭代调整模型，提高人机交互的效率。



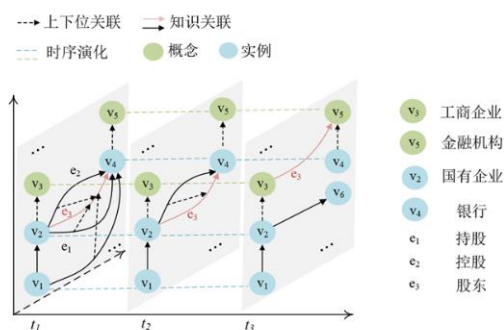
(知意-人在环路知识图谱构建平台)

知融-金融时序知识图谱查询与分析平台

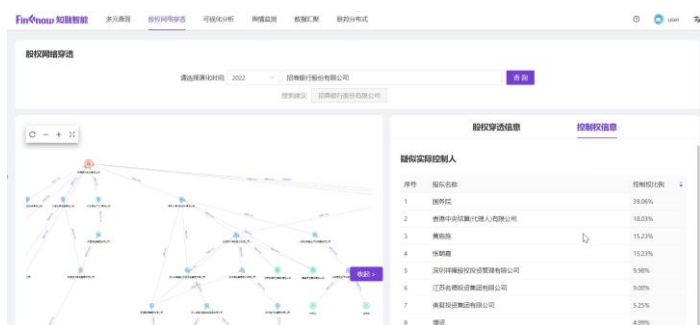
<http://47.104.72.126:9103/user/login>

平台介绍

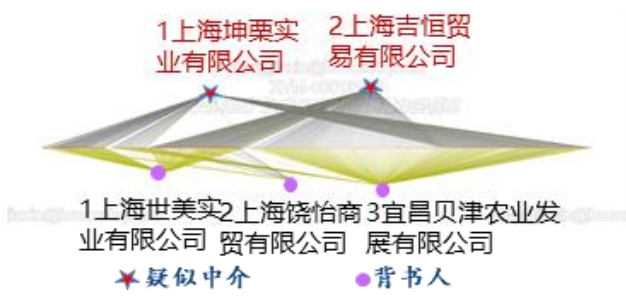
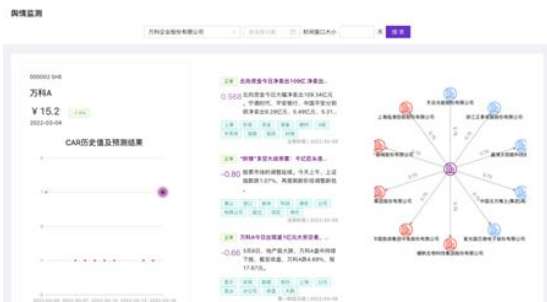
在知意系统基础上，关联融合全量工商注册企业和金融机构股权数据，建立亿级股权知识大图，建设知融-金融时序知识图谱查询与分析平台。对亿级金融知识大图实现了高效的联邦分布式管理。知融平台具有多元查询、股权穿透、舆情监测、控制计算、欺诈识别等功能，形成具备风险发现、风险分析、风险防控全流程能力的金融风控大脑。知融平台不仅有助于风险的发现，还能够进行风险分析，最终实现全流程的风险防控。



▲金融知识大图



▲知识大图股权穿透



(演示视频)

智能决策金融大数据平台

<http://115.156.114.151:8000>

平台介绍

通过整合多源金融数据，包括全量工商注册企业和金融机构股权数据，实现了全面、多维度的数据关联，建立了亿级别的股权知识大图，实现了高效的联邦分布式管理。平台提供多元查询功能，包括股权穿透、舆情监测、控制计算、欺诈识别等，为金融行业提供更全面的决策支持。

