

专题：基于大数据的知识图谱 及其应用

Big Data Based Knowledge Graph and Its Applications

客座编辑



靳小龙 (1976-)，男，博士，中国科学院计算技术研究所研究员、博士生导师，中国科学院网络数据科学与技术重点实验室知识计算方向负责人，大数据分析系统国家工程实验室主任助理；中国科学院大学岗位教授；中国计算机学会大数据专家委员会副秘书长。主要研究方向为知识图谱、知识计算、社会计算、大数据分析等。迄今为止共出版中英文专著5部（分别由Springer与清华大学出版社出版）；在*IEEE TKDE*、*IEEE/ACM TASLP*、*ACM TIST*、*IEEE TWC*、IJCAI、AAAI、WWW、CIKM、WSDM等国际期刊与会议上发表论文200余篇，2019年获得Elsevier Top Cited Articles Award；获得国际会议IEEE CIT-2015、IEEE AINA 2007、IEEE ICAMT 2003的最佳论文奖，IEEE ICBK 2017的最佳学生论文奖，国内会议CCF Big Data 2015的最佳学术论文奖；申请/获得专利20余项。2017年获得国家科学技术进步奖二等奖；2016年获得中国电子学会科技进步奖一等奖。



阮彤 (1973-)，女，博士，华东理工大学教授、博士生导师，华东理工大学计算机技术研究所所长，自然语言处理与大数据挖掘实验室主任。目前兼任中国计算机学会大数据专家委员会委员，中国中文信息学会信息检索专业委员会委员，中国医院协会信息专业委员会委员，上海市经济和信息化委员会公共数据开放专家委员会委员。长期从事自然语言处理与大数据挖掘等方面的科研与应用工作，在大规模文本抽取、知识图谱与数据治理方面获得多项创新成果。

导读

随着大数据时代的到来,知识工程迎来了新的发展机遇。特别是在谷歌公司2012年公布了知识图谱(knowledge graph)项目用于增强其搜索引擎的性能之后,基于大数据的知识图谱迅速得到了学术界、工业界甚至是政府部门的高度关注。近几年,知识图谱已经在搜索引擎、智能问答与自动客服等通用技术,以及电子商务、医疗健康、智慧金融、智能交通、电子政务等应用领域得到广泛且深入的应用。本专题汇集了人工智能与知识工程领域专家学者的8篇文章,一方面,对国内外知识获取、知识推理、知识问答、知识应用以及事件图谱等多个方向的新技术、新方法进行了系统综述;另一方面,为读者梳理了知识图谱在电子商务、金融欺诈、医疗健康等不同领域的应用现状,总结了当下的挑战,展望了发展的趋势。

郑慎鹏等人撰写的《基于主体掩码的实体关系抽取方法》针对当前级联式实体关系抽取技术在主体向量表示和指针网络解码上存在的不足,提出了一种利用注意力和掩码机制生成主体向量并引入实体序列标记任务来辅助指针网络解码实体的新方法。

刘庆霞等人在《实体摘要系统的解释性评测》中,针对现有实体摘要系统以较为复杂的方式集成多种摘要技术特征因而影响了评测结果的可解释性的问题,提出了特征效用率和特征显著率两种指标。将两者进行对比可在一定程度上为摘要系统取得的最终效果提供解释。

在《时态知识图谱补全的方法及其进展》中,申宇铭等人总结了当前基于符号逻辑和基于知识表示学习的时态知识图谱补全方法,对两类方法的优缺点进行了比较和分析。此外,还介绍了用于时态知识图谱补全的基准数据集以及部分模型在其上的性能结果。

王文广在《知识图谱推理:现代的方法与应用》中,用统一的框架系统梳理并介绍了在向量空间中进行知识图谱推理的模型与方

法,将其具体划分为基于欧氏空间和双曲空间的模型与基于各类深度神经网络的模型。

基于知识图谱的问答系统已在电子政务、智慧金融、医疗健康等诸多领域得到广泛应用。在《知识图谱多跳问答推理研究进展、挑战与展望》一文中,社会芳等人对基于知识图谱的多跳问答推理技术的最新研究进展进行了系统综述,介绍了主要的标准数据集和评测指标,最后总结了各类方法模型的优缺点。

事件图谱以事件以及事件之间的关联关系为描述对象的一类特殊的知识图谱。胡志磊等人撰写的《事件图谱的构建、推理与应用》一文,系统概述了事件图谱构建与应用的主流模型及方法,涵盖了事件抽取、事件信息补全、事件关系推断以及未来事件预测等不同方面。此外,介绍了事件图谱的几个代表性应用场景。

陈华钧等人在《大规模知识图谱预训练模型及电商应用》中详细阐述了他们基于十亿规模电商知识图谱所提出的面向不同商品知识服务的“预训练+知识向量”服务模式。实验表明,知识图谱预训练模型可有效地提升商品分类、同款商品识别和商品推荐等电商知识图谱下游任务的性能。

陈强等人的《基于金融知识图谱的会计欺诈风险识别方法研究》将会计案防领域的行业知识与金融知识图谱技术结合,用于精准识别与防范商业银行的会计风险。提出的基于知识图谱的智能化反欺诈方法在银行活期账户的风险排查方面大幅提高了高度可疑账户的识别精准度,同时极大地提升了会计案件的核查效率。

篇幅所限,本专题难以覆盖知识图谱的方方面面,也无法深入探讨各类技术方法的核心细节。但希望通过本专题,让读者对知识图谱的主流技术以及应用现状和前景有一定程度的了解,从而吸引更多人投入这一快速发展的领域之中,共同推动智能化时代的到来。