

学科知识图谱：内涵、技术架构、应用与发展趋势

马富龙, 张泽琳, 闫燕

(西北师范大学教育技术学院, 甘肃兰州 730073)

摘要: 知识图谱作为人工智能技术, 是由感知智能向认知智能发展的关键支撑技术之一, 在教育领域的应用必将推动智能教育的创新发展。首先, 从知识图谱技术发展历程、学科发展历程透视学科知识图谱的技术本质与知识本质, 明确学科知识图谱的价值定位。其次, 在梳理学科知识图谱构建技术与模式演变发展的基础上, 构建学科知识图谱的基本功能架构与技术体系, 分析其在智能教育中的典型应用场景。最后, 从应用场景与技术本体发展两个角度对学科知识图谱的发展趋势进行前景展望, 以期对学科知识图谱助力智能教育应用提供借鉴。

关键词: 学科知识图谱; 技术架构; 发展趋势

DOI: 10.11907/rjdk.232213

开放科学(资源服务)标识码(OSID):



中图分类号: G434

文献标识码: A

文章编号: 1672-7800(2024)003-0212-09

Subject Knowledge Graph: Connotation, Technique Architecture, Application and Development Trend

MA Fulong, ZHANG Zelin, YAN Yan

(College of Educational Technology, Northwest Normal University, Lanzhou 730073, China)

Abstract: As an artificial intelligence technology, knowledge graph is one of the key supporting technologies for the development from perceptual intelligence to cognitive intelligence. Its application in the field of education will undoubtedly promote the innovative development of intelligent education. Firstly, from the perspective of the development process of knowledge graph technology and discipline, the technical and knowledge essence of disciplinary knowledge graph is examined, and the value positioning of disciplinary knowledge graph is clarified. Secondly, based on the analysis of the evolution and development of disciplinary knowledge graph construction technology and models, the basic functional architecture and technical system of disciplinary knowledge graph are constructed, and its typical application scenarios in intelligent education are analyzed. Finally, from the perspectives of application scenarios and technological ontology development, this paper provides a prospect for the development trend of disciplinary knowledge graphs, in order to provide reference for the application of disciplinary knowledge graphs in intelligent education.

Key Words: subject knowledge graph; technique framework; development trend

0 引言

大数据、人工智能等技术在教育领域的创新应用, 促使教育信息化向个性化、智能化与精细化方向发展。以深度学习、知识图谱为核心的新一代人工智能技术的回归, 将对个性化自适应学习系统进行重塑和再造^[1]。《新一代

人工智能发展规划》明确提出, 建立新一代人工智能关键共性技术体系, 研究知识图谱构建与学习技术, 构建覆盖数亿级知识实体的多元、多学科、多数据源的知识图谱^[2]。2021年7月, 《教育部等六部门关于推进教育新型基础设施建设构建高质量教育支撑体系的指导意见》明确指出, 在数字资源新型基础设施重点方向上要系统梳理各学科知识脉络, 明确各知识点间的关系, 分步构建国家统一的

收稿日期: 2023-11-29

基金项目: 国家自然科学基金项目(62167007); 甘肃省自然科学基金项目(22JR5RA157)

作者简介: 马富龙(1995-), 男, 西北师范大学教育技术学院硕士研究生, 研究方向为教师专业发展; 张泽琳(1997-), 女, 西北师范大学教育技术学院硕士研究生, 研究方向为中小学信息科技教育; 闫燕(1996-), 女, 西北师范大学教育技术学院博士研究生, 研究方向为中小学信息科技教育。本文通讯作者: 张泽琳。

学科知识图谱,并对现有资源进行分类标识,使其匹配学科知识图谱^[3]。

在此背景下,探究以知识图谱为代表的新一代智能技术在教育领域中的应用成为推进智能教育发展的重要议题。鉴于此,本文梳理总结学科知识图谱发展历程,明晰学科知识图谱的内涵、技术体系及发展趋势,为学科知识图谱在教育场景中的应用提供参考与借鉴。

1 学科知识图谱的内涵

1.1 从知识图谱技术发展历程透视学科知识图谱的技术本质

1.1.1 发展历程

知识图谱是知识工程在当前人工智能、大数据技术快速发展时期的新形态与新产物,现在正与其他新兴信息交叉融合发展。中国中文信息学会语言与知识计算专委会对知识工程四十年以来发展历程进行了回顾,发现知识工程由 5 个标志性的阶段组成:前知识工程时期、专家系统时期、万维网 1.0 时期、群体智能时期及知识图谱时期^[4]。其中,知识图谱发展历程有几个关键事件与概念,可以此为关键特征对知识图谱的技术发展历程进行划分,如图 1 所示。

(1)专家系统概念的提出——知识图谱技术的萌芽时期。人工智能的符号主义学派认为人的认知过程是在符号表示上的一种运算^[5],倡导计算机模拟人类思维活动,通过逻辑表达方式实现计算机智能。1965 年,斯坦福大学的 Edward Feigenbaum 提出专家系统 (Expert System) 的概念,原理是在领域知识库基础上通过推理机制解决问题^[6],此时的知识库是知识图谱的最早雏形。

(2)知识工程概念的提出——知识图谱技术的工程化时期。1977 年,美国斯坦福大学计算机科学家 Edward Feigenbaum 在举办的第五届国际人工智能会议上提出知识工程的概念,希望将特定领域专家的知识集成到计算机系

统,让计算机系统依据知识库、知识规则完成领域专家才能完成的推理、问答等复杂任务,进而辅助普通用户能以领域专家的水准完成任务。此外,知识工程概念的提出使人们从系统化的视角思考与建设知识图谱。

(3)知识本体概念的提出——知识图谱技术的形式化时期。1991 年,Niches 提出知识本体的概念将领域知识形式化。1993 年,Gruber 定义知识本体 (ontology) 是概念体系的明确规范 (An ontology is an explicit specification of conceptualization)。研究者通过捕获相关领域的知识并从不同层次对其进行规范化描述,使其具有明确定义、可被计算机处理、能实现群体共享。知识本体概念的提出与应用,为知识图谱技术从理论研究走向大规模工程实践奠定了基础。

(4)语义网与专家知识系统发展——知识图谱技术的互联网时期。1998 年,万维网 (World WideWeb, WWW) 的发明者 Tim Berners-Lee 在万维网发展 10 年后提出语义网 (Semantic Web, SW) 概念,期望将传统人工智能的发展与万维网结合,以资源描述框架 (Resource Description Framework, RDF) 为基础在万维网中应用知识表示与推理方法^[7],因此出现了大规模的专家知识系统。例如,第一代专家系统 DENRAL 识别分子结构系统、第二代专家系统 MYCIN 医疗诊断系统等。此时,知识图谱技术开始与互联网技术结合开展应用,知识图谱走向网络化。

(5)知识图谱概念提出——知识图谱技术的快速发展时期。2012 年,Google 正式提出知识图谱 (Knowledge Graph) 概念,构建了用于增强搜索引擎功能的知识库^[8],开启了现代知识图谱序章。此时,知识图谱在搜索引擎、电商、医疗、安防、金融、教育等领域开始应用,例如 Facebook、百度、阿里巴巴、腾讯等互联网企业构建的商业领域知识图谱及 eduKG、MOOCube 等典型的教育领域知识图谱等。其中,知识图谱与其他技术融合发展,尤其是人工智能、大数据技术,使其构建走向半自动化、自动化阶段,在智能系统的作用逐步向上支持各类复杂服务,向下统摄各类数据。

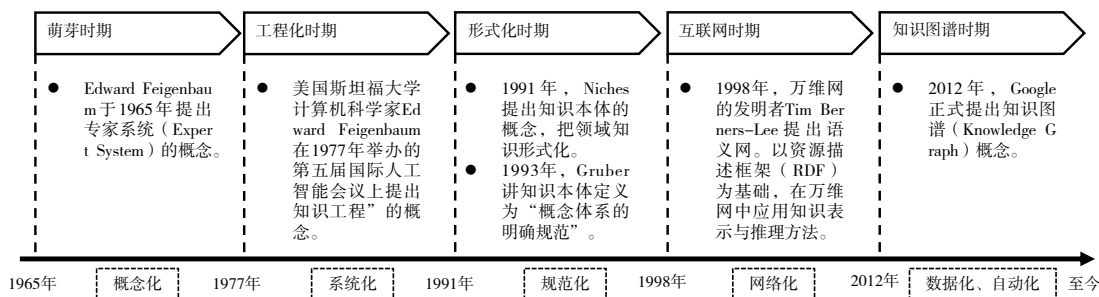


Fig. 1 Development, key concepts and characteristics of knowledge graph technology

图 1 知识图谱技术发展历程、关键概念与特征

1.1.2 学科知识图谱的技术本质

纵观知识图谱发展历程与特征,在与各类技术不断融合中得到发展并努力实现自身价值最大化。从技术视角

而言,知识图谱涉及知识抽取与表示、数据存储与查询、知识融合和知识推理等方面,因此在其构建与应用的整个生命周期中继承了知识工程、互联网、语义网、自然语言处

理、数据库、人工智能等技术基因,形成了以图谱数据为基础与核心,集成算法、工具、应用系统的技术统一体。从语义网角度而言,知识图谱为数据构建了类似语义网文本之间的超链接语义链接,并支持语义检索。从自然语言处理角度而言,知识图谱的构建需要从非结构化的文本中抽取语义和结构化数据。从知识表示角度而言,构建知识图谱需要通过计算机符号表示和处理知识。从人工智能角度而言,构建知识图谱需要利用知识库辅助理解人类语言。从数据库角度而言,构建知识图谱需要采用某种方式存储知识。

此外,学科知识图谱作为学科教学领域的图谱,虽然有别于通用知识图谱,但从技术本质而言一致,是获取与表示知识,并提供知识服务的技术统一体,如图2所示。

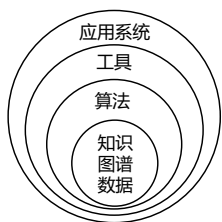


Fig. 2 Technical essence of subject knowledge graph

图2 学科知识图谱技术本质

1.2 从学科发展视角透视学科知识图谱的知识本质

1.2.1 发展历程

学科是知识分类的产物,是一定知识领域或一门科学的分支,是人类知识体系的基本单元。知识是人们在改造世界的实践中所获得的认识和经验的总和^[9],其发展经历了从零散知识到体系化知识系统的演变过程,将知识按照其内在联系或相类似的表现形式加以划分,形成的一个个知识集合即为学科。

早期,人类为了生存学会分类植物、动物和无生命体等,并且简单掌握了这些事物之间的关系,形成最初的知识^[10]。由于社会发展、人们对知识的认识不断提高、知识运用频率提升、知识的快速增长,人们开始对知识进行总结与分类,采用不同分类标准对知识进行划分,逐步形成了学科。例如,孔子对以往的典籍进行编排整理,选择六艺开展分科教学;培根依据近代科学知识的科学性、普适性、规律性与客观性等特点划分科学知识体系,绘出有巨大影响的“知识树”图解等^[11]。学科存在科学研究中的学科和教育中的学科^[12]。在教育领域,学校教育研究者从教育的视角出发,从语言、社会、自然、艺术、体育等文化领域和生活、道德等活动领域的知识中选择内容,从心理学视角加以系统化编制,形成了语文、数学、英语、体育等科目,在实践中表现为教材结构。

1.2.2 知识本质

从学科发展历程而言,学科将人类社会发展历程中获得的认识和经验建立为知识体系,而学科知识图谱则将知识体系进行图式表征,基于此实现大规模的高效关系运

算、复杂关系推理、潜在关系挖掘等。由此可见,学科知识图谱具有知识本质,将学科概念和符号、原理及格式、过程方法、认知策略、事实范例和价值观等各类知识以符号形式进行描述,通过语义实现知识节点之间的连接,形成知识网络,借助图的查询、追溯、可视化等技术,实现知识间复杂关系网络的发现、理解、推理与分析。

(1)学科知识图谱边界。通常按照知识概念的范畴对学科知识图谱的外延进行界定:广义的学科知识图谱是将学科知识进行图谱化表征的结果;狭义的学科知识图谱是将学校教育中教学科目的具体知识进行图谱化表征的结果。具体关系如图3所示。



Fig. 3 Relationship among course knowledge graph, subject knowledge graph and knowledge graph

图3 科目知识图谱、学科知识图谱、知识图谱之间的关系

(2)学科知识图谱特征。学科知识图谱的知识本质使其自然继承知识的特征:①学科知识图谱具有科学性特征,由于学科知识是对客观事物发展规律的客观总结和概括,因此学科知识体系客观存在科学性特征^[12],这也是学科知识图谱最基础的特性;②学科知识图谱具有系统性特征,学科是对领域知识进行体系化总结的产物,自身具有蕴含逻辑结构的概念关系网,在细分领域是一个完整的系统化表达;③学科知识图谱具有动态变化特征,由于学科领域研究者不断探索和实践,尤其是自然科学领域研究者对具体知识的不断实践和深挖,学科知识体系一直处于动态变化中,需要将新知识补充到原有学科知识体系;④学科知识图谱具有特殊性,知识发展为学科的必要条件是学科具有独有的概念体系、表达方式和研究方法,因此在对知识进行图谱化表征时,抽取的知识节点在性质上属于该学科特有的中心概念,知识的表征要遵循学科独特的表达方式、技术和技巧;⑤学科知识图谱具有复杂性特征,随着信息技术发展,学科知识在呈现方式上走向多样化,例如文字、图片、音视频、动画、AR/VR等形式,但也增加了学科知识图谱的构建与应用场景的复杂度。

1.3 学科知识图谱是学科知识和图谱技术在知识价值实现目标下的统一

目前,信息社会的领域分工与合作日益加深,一方面各领域研究分工越来越细化,知识增速提升,对知识细分的要求越来越高,领域研究者迫切需要对知识进行体系化表征,构建细分领域的知识框架,为自身研究及后续研究者提供支持;另一方面,研究和实践逐步走向跨领域合作,知识发展具有综合化、整体化的趋势,亟需打破现有知识分类(学科是知识分类的结果)的界限,建立各领域知识间的普遍性联系,为创新研究提供支持。

在此背景下,学科知识图谱的构建恰逢其时,既提供了体系化知识表征,又建立了知识间的普遍联系,为信息

社会知识价值实现最大化提供支撑,如图 4 所示。

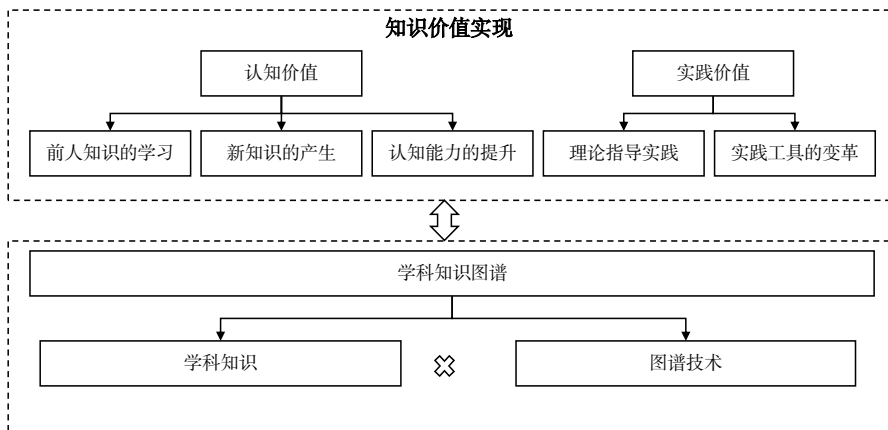


Fig. 4 Positioning of subject knowledge graph

图 4 学科知识图谱定位

1.3.1 知识的价值

价值是客体及其属性能满足主体需要的意义。知识的价值是价值的一种表现形式,是知识客体满足人类主体需要的意义,主要表现在其认知价值与实践价值两个方面。

(1) 认识价值。知识能使我们接受前人和他人的认识成果,作为进一步认识事物的基础。新的认识成果是人类运用恰当的思维方法,对已积累知识加以延伸和碰撞而产生,因此只有积累了对某个事物的大量知识,建立了暂时联系,才能在新事物出现时通过直觉思维或者灵感快速作出识别和判断。其次,知识是人类认识事物思维能力提升的重要条件,认识事物的目标是发现规律,而规律的获得需要经历提出问题和假设、验证假设到作出结论的思维过程,只有对该事物的知识积累越广、越深,在思维过程中分析问题和解决问题时思路才能更灵活、方法更多样、判断更精准,思维能力才能得到锻炼与提升。

(2) 知识的实践价值。首先,知识能够指导实践活动。理论知识反映了客观事物的规律,利用科学理论知识指导实践达到改造世界的目的,就是知识与实践结合的过程。其次,知识有助于提升人类实践技能和水平。在人类实践过程中,通过知识成果的转化推动技术的重大改进,不断变革生产工具、手段、方法进而提升实践效率。

1.3.2 知识图谱在知识价值实现中的基础支撑作用

在实现知识认知价值的视野下,需要将知识体系化表征出来,提供知识的检索、学习等基础服务,便于认知主体、了解和掌握知识。知识图谱支持的学习系统能为学习者提供一种体系化的知识呈现方式和学习路径,辅助学习者构建学科领域知识体系,因此需要表征领域内外知识间的联系,便于主体在认知过程中提供推理服务,辅助主体的高阶认知。综上,基于知识图谱的自适应学习系统能为学习者构建一个高阶思维能力培养的环境,辅助、引导学习者使用关联思维思考和解决问题。

在实现知识实践价值的视域下,实践需要在人类积累的大量知识指导下开展,避免走弯路、走错路,知识图谱正是对这些庞大、复杂的理论和经验的体系化表征,为指导实践提供知识库服务。例如,用户在洗衣机使用过程中出现问题,能在商业平台中智能客服机器人的指导下排除故障,该场景中的智能客服机器人就是在有关洗衣机知识库的支持下对用户的问题进行诊断和推理,指导用户解决问题。此外,通过知识的转化推动实践工具能力和实践效率的提升,例如医学领域借助知识图谱的影像分析系统辅助病情的诊断提升诊断效率。

2 学科知识图谱技术体系与典型应用

2.1 学科知识图谱构建技术与模式演变

知识图谱构建技术的发展推动构建模式的演变,知识图谱构建包括知识表示与抽取、知识融合、知识推理等方面的技术^[13]。随着构建技术发展,学科知识图谱已从单一技术、以人类专家为主的构建方式走向系统工程、人机协同构建方式,大致经历了传统人工式构建、半自动化的机器构建和基于深度学习的智能化3个阶段,如表 1 所示。

2.2 学科知识图谱的基本架构

知识图谱构建技术与模式的演变不断完善与细化知识图谱基本架构,从二层架构逐步发展至四层架构。早期,知识图谱以人工构建为主,从逻辑上将其划分为数据层、模式层,数据层采用三元组形式存储知识,模式层采用本体库规范实体、关系、属性及其之间的联系;随着知识抽取与加工技术不断发展,知识图谱架构发展为知识获取、知识融合层、知识计算层与应用层,知识获取层从多源异构的数据源中提取出实体、实体关系及其属性等知识要素,知识融合层实现多知识的多数据源关联、实体对齐、属性融合,知识计算与应用层实现对知识图谱服务能力的输出;近年来,知识图谱构建技术进一步细分,各层之间的任

Table 1 Evolution of knowledge graph construction techniques and modes

表1 知识图谱构建技术与模式演变

	传统人工式构建	半自动化的机器构建	智能化构建
知识表示与抽取	采用人工编写规则与模板的方式对实体进行抽取	采用基于统计机器学习的方式,对数据进行标注与模型训练,例如KNN、CRF等统计分析模型用于命名实体与关系的抽取 ^[14]	采用卷积神经网络、循环神经网络模型进行实体抽取 ^[15,16] 基于参数共享和基于序列标注的联合抽取方法 ^[17]
知识融合	人工判断方式	多采用概率模型和机器学习的方法实现实体对齐,逐步将匹配问题转换为分类问题进行对齐	采用TF-IDF进行局部对齐 ^[18] 采用关系聚类、SiGMa算法、CRF及Markov逻辑网模型实现全局实体对齐 ^[19]
知识推理	基于逻辑规则的推理技术	基于分布式特征的推理技术	基于神经网络的深度学习推理技术

务界限更清晰,逐步发展为数据处理、知识加工、知识管理及知识服务4层,本文依据对学科知识图谱内涵的分析及技术的发展变化,提出知识获取、知识融合、知识存储、知识引擎与知识应用服务5层架构,如图5所示。

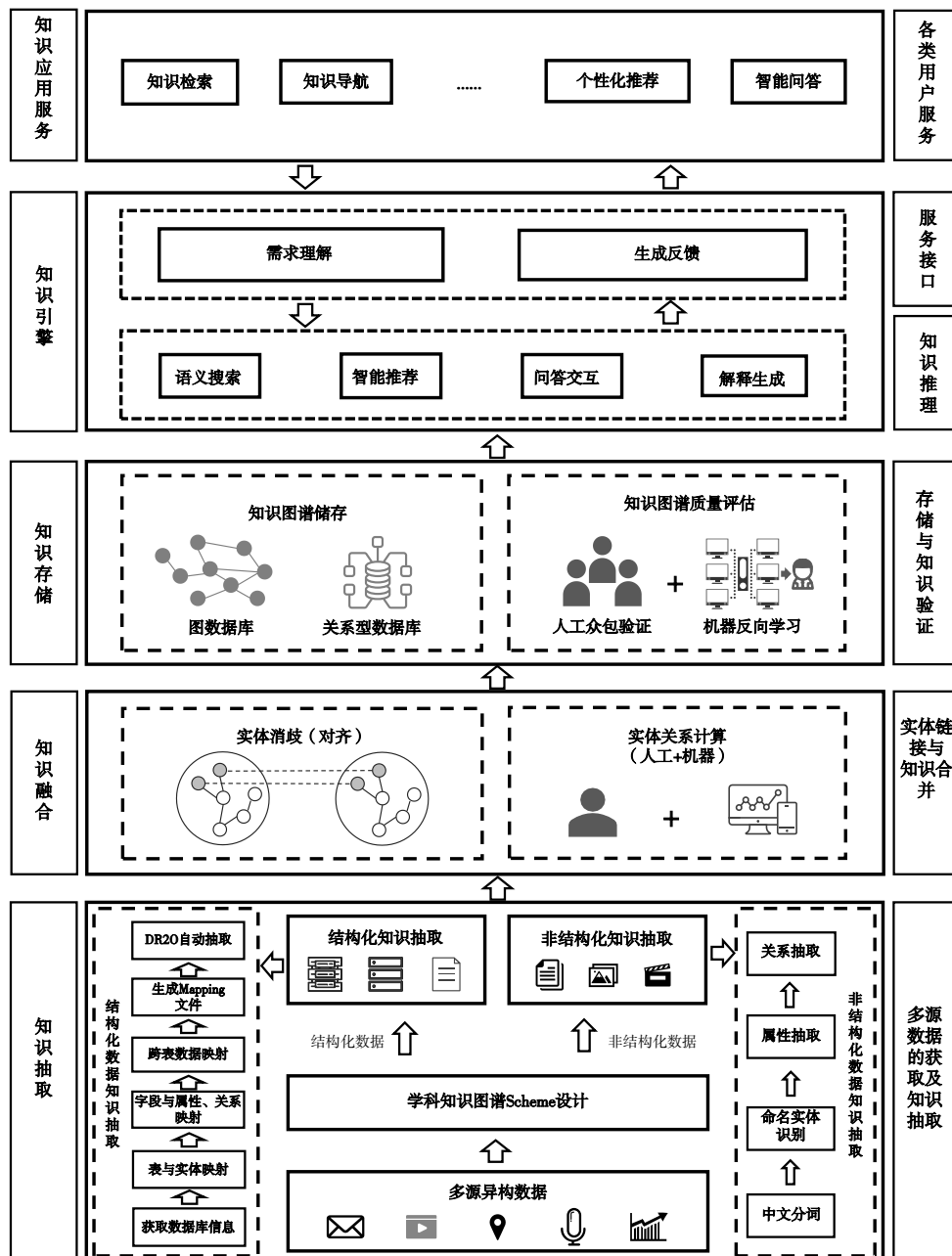


Fig. 5 5-layer architecture of subject knowledge graph

图5 学科知识图谱的5层架构

(1)知识抽取层。首先,采集多源异构数据,包括课程标准、教材、教辅、在线百科、核心词汇、试题集等教育领域数据;DBPedia、网页表格与数据等开放领域数据。然后,建立知识表示模式,对采集的结构化数据与非结构化数据进行知识抽取,形成知识实体。

(2)知识融合层。针对各类知识实体进行属性、关系及实例等方面的计算,通过实体链接、本体扩充等时序融合与实体匹配、概念对齐等多源融合相结合的方式实现学科知识的融合,并通过学科专家进行知识验证与确认,确保所构建的学科知识图谱的科学性。

(3)知识存储层。主要将构建的学科知识及其关系以图数据库或关系数据库等方式进行存储,为知识的应用与可视化提供支撑。同时,基于存储的知识图谱提供人工众

包验证的工具与机器反向学习数据集,实现知识图谱质量的评估。

(4)知识引擎层。该层接受应用层的需求,通过自然语言理解识别用户需求,并将其转化为机器指令与知识图谱进行交互,通过语义搜索、智能推荐、问答交互、解释生成等智能处理过程,产生反馈信息并推送至知识应用服务层。

(5)知识应用服务层。学科知识图谱根据教育应用场景的需求,以嵌入或独立方式向各类应用系统提供知识检索、知识导航、个性化推荐、智能问答等服务^[20-21]。

2.3 学科知识图谱的技术体系与各层关键技术

依据学科知识图谱的五层架构,梳理构建对应的技术体系,如图6所示。

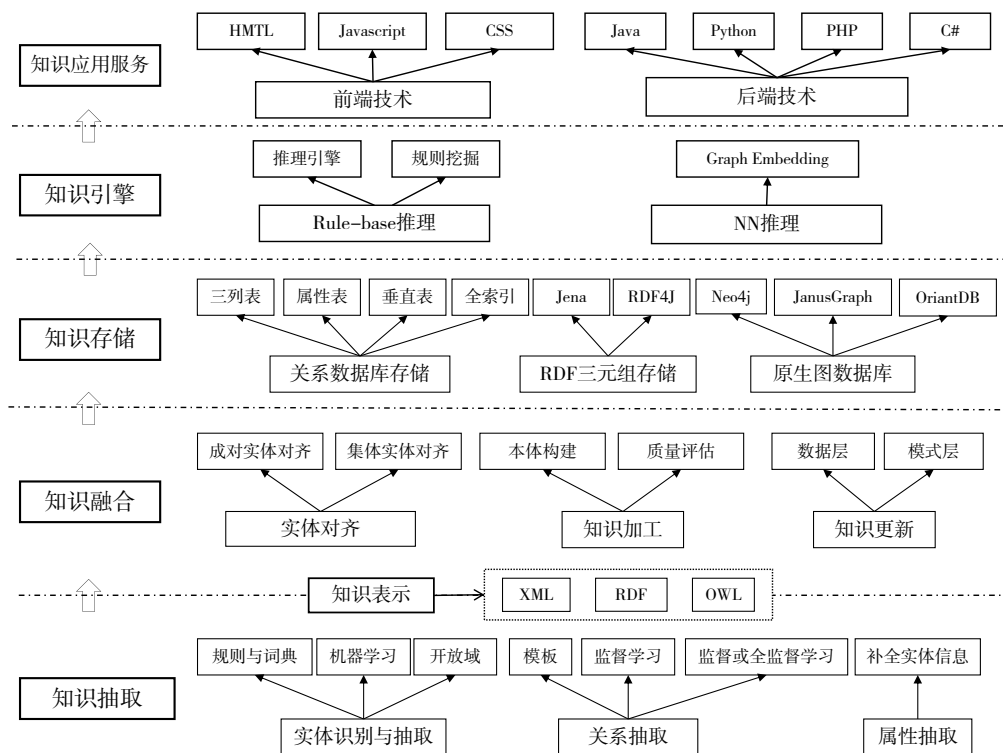


Fig. 6 Technical system of subject knowledge graph

图6 学科知识图谱的技术体系

2.3.1 知识抽取层

知识抽取主要面向开放的链接数据,关键技术是自然语言处理(Natural Language Processing, NLP)技术,以实体抽取,也叫命名实体识别(Named-entity recognition, NER)与实体链接、关系抽取、属性抽取为主,为知识融合奠定基础^[22]。

(1)实体抽取。识别命名实体的文本范围,并将其分类为预定义的类别,主要包括基于规则与词典、基于统计机器学习及面向开放域的方法。其中,基于规则与词典方法通常选用规则模版,利用模式和字符串相匹配为主要手段进行匹配;基于机器学习方法将命名实体识别、分词问题作为序列标注问题,预测标签之间具有强的相互依赖关

系;面向开放域的抽取是面向海量的Web语料。

(2)关系抽取。关系抽取目标是建立实体间语义链接,包括基于模板、基于监督学习、基于半监督或无监督学习3种抽取方式^[23]。其中,基于模板通过人工或机器学习方法抽取实体关系适用于小规模、特定区域内的场景;基于监督的关系一般将抽取问题转化为分类问题,模型的选择主要包含支持向量机、朴素贝叶斯等机器学习分类模型;基于半监督或无监督学习的关系抽取通过知识图谱与非结构化文本对齐的方式自动构建大量数据集,减少模型对人工标注数据的依赖。

(3)属性抽取。实行抽取要识别实体的实行名和属性值,因此大多基于规则进行抽取,面向的对象也是网页、

query、表格数据^[24]。

综上,知识表示是对知识抽取后信息进行有效组织和表述的关键技术,通常使用资源描述框架(Resource Description Framework, RDF)、网络本体语言(Web Ontology Language, OWL)对知识进行描述,这两者都使用可扩展标记语言(Extensible Markup Language, XML)作为核心语法。

2.3.2 知识融合层

知识融合层主要整合不同数据源的数据,形成高质量的知识,包括实体对齐、知识加工与更新技术。

(1)实体对齐。实体对齐也称为实体消解、实例匹配,可分为成对实体对齐与集体实体对齐两大类,主要为了消解知识图谱融合过程中的冲突,对知识进行关联与合并,最终形成一致的结果。

(2)知识加工。知识加工通过本体构建、知识推理和质量评估3个阶段获得结构化、网络化的知识体系。

(3)知识更新。通常以自动方式更新本体、实体中的元素,包括概念、概念属性、概念之间上下位关系、实体、实体属性值等。

2.3.3 知识存储层

依据知识图谱的规模、图谱应用场景、应用方式等,在实际生产中图谱存储往往采用不同的技术路线。

(1)关系数据库存储。基于关系表对知识图谱数据进行组织,主要包括基于三列表的存储方式、属性表的存储方式、垂直表的存储方式及全索引的存储方式等,实体之间的关系通常通过外键实现,适用于图谱规模不大且实体关系简单的场景。

(2)RDF三元组数据库存储。主要适用于存储RDF数据开发的图谱数据库,支持SPARQL查询语言,开源的RDF三元组数据库包括Jena、RDF4J、RDF-3X和gStore等。

(3)原生图数据库存储。一种非关系型数据库指的是以图的方式存储、处理、查询和展现数据,能较好地解决有关系数据库的局限性。图的特性能较好地展现数据间的依赖关系,其他数据库模型则通过隐式连接链接数据,并简单快速的检索建模中的复杂层次关系。目前,典型的图数据库有Neo4j、JanusGraph、OrientDB等。

2.3.4 知识引擎层

知识图谱引擎基于高质量知识图谱数据,实现知识推理核心业务逻辑并向上提供服务接口的集成化平台。通过NLP技术、基于规则的推理等算法识别、理解用户需求,然后基于对知识图谱数据的高速检索、知识追踪、逻辑推理等生成反馈数据。

2.3.5 知识应用服务层

该层利用高级编程语言,基于知识引擎所提供的服务接口,构建各类场景中的应用服务,并通过可视化的人机交互界面提供给应用服务用户,包括前端HTML、CSS、JavaScript等技术及Java、C#、PHP、Python等主流后端技术。

在实际工程化过程中,由于学科知识图谱规模、构建模式与方法、应用场景与方式等方面存在差异,因此基本架构可能会进一步细分或合并。随着技术不断发展,各层的关键技术也可能进一步细化或与其他层级的技术实现交叉融合,因此研究与实践者应正确认知并及时运用新技术进行优化。

2.4 学科知识图谱的典型应用

(1)用户画像。随着移动互联网的飞速发展,教育工作者越来越多地应用现代数字技术开展教育教学,师生在教学和学习过程中时刻产生各类数据,例如平台登录基本信息、学习风格、教学风格、资源设计开发、分享记录等,这些数据具有重要价值,可为资源推荐、学生和教师发展、教学管理等提供依据,而用户画像又服务于这些应用。研究者构建了基于学科知识图谱的学习者画像模型^[25],通过分析学生的学习行为、学习数据,了解学生学习状态、判断学生能力达成,从而为学生规划个性化学习路径^[26]。此外,基于学科知识图谱的师生画像建模能提升教学过程中数据的可解释性,可更好地支持学生的个性化发展^[27]。

(2)智能问答系统。知识图谱逐渐应用于各行各业,人们开始关注基于知识图谱的智能问答系统,例如医疗^[28]、旅游^[29]、农业^[30]、电商^[31]等行业的智能问答系统正在研究及应用。在教育领域,学科知识图谱将知识点进行分类、关联和解析,建立智能问答系统,帮助学生解决学习中的问题,从而深入理解和掌握知识点。例如,部分研究者基于学科知识图谱构建英语知识问答系统^[32]、设计开发育人助理系统、实现个性化辅导^[33]等。

(3)资源精准推荐。基于知识图谱的语义理解,可为学生提供个性化的学习资源推荐,提升学习效率。知识图谱的个性化推荐模型相关研究表明^[34],实时诊断学生学习情况,结合学习者个人需求为学生系统推荐学习路径、同伴、资源、习题等^[35]意义重大,以学科能力为基础的“智能学伴”可融合学科知识图谱,支持面向学生学科能力的自主学习模式和课堂教学模式^[36]。因此,恰当匹配教育资源与个性化学习服务,促进教育公平实现和优质教育资源共享是学科知识图谱赋能智能教学的重要应用^[37]。

3 学科知识图谱的发展趋势

在知识图谱技术本体发展与教育应用场景需求的双向驱动下,知识图谱与多模态技术、事理图谱、区块链技术动态融合与协同发展,学科知识图谱的技术体系将不断完善发展,对教育的关键支撑作用将进一步得到发挥。

3.1 与多模态技术融合推动资源生态体系构建

多模态融合技术不断发展,为数据集的跨模态迁移学习提供了桥梁;多模态数据的识别与分析能力大幅增强^[38],提升了学科知识图谱链接多模态数据的能力。教育资源生态体系的构建,迫切需要构建资源本身内在关联体

系,也需要与社会资源体系建立关联。

在学科知识图谱中融入多模态技术,既能有效标注、关联教育领域内资源,还将博物馆、研究机构、社会领域的资源有机纳入资源体系,以多形态表征为学习者构建立体化资源体系,为学习者提供多样化、个性化的资源奠定基础^[39]。随着机器学习、数据挖掘及自然语言处理等技术发展,学科知识图谱融合多模态教育资源成为学科知识图谱未来发展趋势之一^[20]。

3.2 与事理图谱融合推动教育智能体发展

事理知识图谱将事件间的因果、顺承、条件、互反等逻辑关系抽取并抽象出来,以图谱形式描述事件间的演化规律和模式提升图谱推理能力。智能教育的发展,对教育机器人、智能代理、智能学伴等教育智能体的推理、问答、分析能力提出更高要求,将事理知识库与学科知识图谱融合,作为教育智能体的底层支撑技术,将在提升智能体智能诊断能力的基础上大幅提升其在知识问答、智能推送、对话生成、辅助决策等方面的能力,为学习者提供智能辅导、自适应学习、智能陪伴等个性化学习服务。

3.3 与区块链技术融合推动学科知识共享与生成

区块链技术的可追溯、可信、公开、共享等特性,为知识图谱的数据存储、分布式计算、知识确权等提供支撑。随着学科知识体系的动态变化,学科内部与学科间的知识融合、构建完整和科学的知识体系的要求不断提升,知识创新与生成模式的走向具有扁平化、群体化等特征。

在学习知识图谱构建中融入区块链技术,发挥其在知识确权与保护、知识交易上的优势,推动知识共享新生态的构建,建立学科知识图谱网络众包机制,凝聚互联网智力资源,协调发包方、接包方、平台方及其他利益相关者在知识生成过程中的利益,实现“网络众筹+机器智能”的知识生成模式,推进学科知识图谱的动态迭代更新。

4 结语

学科知识图谱的建设与应用已得到教育管理部门、科研工作者和教育产品开发者的高度重视,已呈现百花齐放、百家争鸣的势头,应正确认识学科知识图谱的建设与应用是一个系统工程,无法一蹴而就。

在此关键时期,应该科学认知学科知识图谱内涵与本质,将为顶层设计者尊重规律、系统谋划提供方向指引;科学认知功能架构及技术体系,将为研究与实践者实现关键技术突破与产品研发提供路线图;科学认知未来发展趋势,将为各利益相关者提供目标引领与结果导向。为此,本文从上述3个维度进行初探,以期系统化研究与实践学科知识图谱的学者提供参考与借鉴。

参考文献:

- [1] MOU Z J. Rethinking and solving personalized learning theory in the era of "artificial intelligence +" [J]. Journal of Distance Education, 2017, 35(2):22-30.

牟智佳.“人工智能+”时代的个性化学习理论重思与开解[J]. 远程教育杂志,2017,35(2):22-30.

- [2] State Council. State Council on the notice issued by a new generation of artificial intelligence development planning [EB/OL]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- 国务院. 国务院关于印发新一代人工智能发展规划的通知[EB/OL]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- [3] Ministry of Education. Ministry of Education and so on six departments of promoting education of new infrastructure building a supporting system for the quality education guidance [EB/OL]. http://www.moe.gov.cn/srcsite/A16/s3342/202107/t20210720_545783.html.
- 教育部. 教育部等六部门关于推进教育新型基础设施建设构建高质量教育支撑体系的指导意见[EB/OL]. http://www.moe.gov.cn/srcsite/A16/s3342/202107/t20210720_545783.html.
- [4] Chinese Language Information Society, Special Committee on Language and Knowledge Computing. Introduction to knowledge graph development report (2018) [EB/OL]. https://wenku.baidu.com/view/c25fe0a5bad528ea81c758f5f61fb7360b4c2b64.html?_wks_=1708395121652.
- 中国中文信息学会,语言与知识计算专委会. 知识图谱发展报告(2018)[EB/OL]. https://wenku.baidu.com/view/c25fe0a5bad528ea81c758f5f61fb7360b4c2b64.html?_wks_=1708395121652.
- [5] ZOU L. Knowledge mapping principle and the application overview [EB/OL]. <https://k.cnki.net/CInfo/Index/2938>.
- 邹磊. 知识图谱原理和应用概述[EB/OL]. <https://k.cnki.net/CInfo/Index/2938>.
- [6] QUINLAN J R. An introduction to knowledge-based expert systems [J]. Australian Computer Journal, 1980, 12:56-62.
- [7] TIAN L, ZHANG J C, ZHANG J H, et al. Knowledge graph: representation, construction, reasoning and knowledge hypergraph theory [J]. Journal of Computer Applications, 2021, 41(8):2161-2186.
- 田玲,张谨川,张晋豪,等. 知识图谱综述——表示、构建、推理与知识超图理论[J]. 计算机应用, 2021, 41(8):2161-2186.
- [8] SINGHAL A. Introducing the knowledge graph: things, not strings [EB/OL]. <https://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html>.
- [9] WIKIPEDIA. Knowledge graph [EB/OL]. https://en.wikipedia.org/wiki/Knowledge_Graph.
- [10] Dictionary Editing Room, Institute of Linguistics, Chinese Academy of Social Sciences. Modern Chinese dictionary (2002 Supplement) [M]. Beijing: Beijing Commercial Printing Museum, 2002.
- 中国社会科学院语言研究所词典编辑室. 现代汉语词典(2002年增补本)[M]. 北京:北京商务印刷馆,2002.
- [11] CHEN H L. Knowledge classification and knowledge resource epistemology [M]. Beijing: People's Publishing House, 2008.
- 陈洪澜. 知识分类与知识资源认识论[M]. 北京:人民出版社,2008.
- [12] WEI D Y. Reunderstanding of subject essence —— from the perspective of discipline history [J]. Journal of Yangzhou University (Higher Education Research Edition), 2015, 19(1):10-12.
- 韦冬余. 学科本质的再认识——学科史的视角[J]. 扬州大学学报(高教研究版), 2015, 19(1):10-12.
- [13] MA Z G, NI R Y, YU K H. The latest progress, key technologies and challenges of knowledge graph [J]. Chinese Journal of Engineering Science, 2020, 42(10):1254-1266.
- 马忠贵,倪润宇,余开航. 知识图谱的最新进展、关键技术和挑战[J]. 工程科学学报, 2020, 42(10):1254-1266.
- [14] RAU L F. Extracting company names from text [C]// 7th IEEE Conference on Artificial Intelligence Application, 1991:29-32.

- [15] RONAN C, JASON W, LÉON B, et al. Natural language processing(almost) from scratch[J]. Journal of Machine Learning Research, 2011, 12: 2493-2537.
- [16] HUANG Z, WEI X, KAI Y. Bidirectional LSTM-CRF models for sequence tagging[DB/OL]. <https://arxiv.org/abs/1508.01991>.
- [17] ZHANG J X, ZHANG X S, WU C X, et al. Overview of knowledge graph construction technology[J]. Computer Engineering, 2022, 48(3): 23-37.
张吉祥, 张祥森, 武长旭, 等. 知识图谱构建技术综述[J]. 计算机工程, 2022, 48(3): 23-37.
- [18] LI J, WANG Z, XIAO Z, et al. Large scale instance matching via multiple indexes and candidate selection [J]. Knowledge-Based Systems, 2013, 50: 112-120.
- [19] BHATTACHARYA I, GETOOR L. Collective entity resolution in relational data [J]. ACM Transactions on Knowledge Discovery from Data, 2007, 1(1): 5-es.
- [20] GAO M, ZHANG L P. Research on the connotation, technology, and application of educational knowledge graph integrating multimodal resources [J]. Computer Application Research, 2022, 39(8): 2257-2267.
高茂, 张丽萍. 融合多模态资源的教育知识图谱的内涵、技术与应用研究[J]. 计算机应用研究, 2022, 39(8): 2257-2267.
- [21] SHEN H Y, XIAO W, JI Y M, et al. Types, applications and challenges of educational knowledge graph [J]. Software Guide, 2023, 22(10): 237-243.
沈红叶, 肖婉, 季一木, 等. 教育知识图谱的类型、应用及挑战[J]. 软件导刊, 2023, 22(10): 237-243.
- [22] WEN Q H, ZHU H Y, HOU L, et al. Multi strategy Chinese open relationship extraction method [J]. Chinese Journal of Information Science, 2023, 37(1): 88-96.
温清华, 朱洪银, 侯磊, 等. 多策略中文开放关系抽取方法[J]. 中文信息学报, 2023, 37(1): 88-96.
- [23] YUAN Q B, DU X M, YANG F. Review of research on limited domain relation extraction [J]. Computer System Applications, 2021, 30(9): 24-40.
袁清波, 杜晓明, 杨帆. 限定域关系抽取研究综述[J]. 计算机系统应用, 2021, 30(9): 24-40.
- [24] XU Q T, HONG Y, PAN Y C, et al. A review of attribute extraction research[J]. Journal of Software, 2023, 34(2): 690-711.
徐庆婷, 洪宇, 潘雨晨, 等. 属性抽取研究综述[J]. 软件学报, 2023, 34(2): 690-711.
- [25] ZHAO L L, FAN J R, ZHAO Y T, et al. Design and application of learner portrait model based on knowledge graph: a case study of high school physics course [J]. Modern Educational Technology, 2021, 31(2): 95-101.
赵玲朗, 范佳荣, 赵一婷, 等. 基于知识图谱的学习者画像模型设计与应用——以“高中物理”课程为例[J]. 现代教育技术, 2021, 31(2): 95-101.
- [26] TANG Y W, RU L N, FAN J R, et al. Research on personalized learning path planning based on learner portrait modeling [J]. E-education Research, 2019, 40(10): 53-60.
唐烨伟, 茹丽娜, 范佳荣, 等. 基于学习者画像建模的个性化学习路径规划研究[J]. 电化教育研究, 2019, 40(10): 53-60.
- [27] LI Z, ZHOU D D, WANG Y. Research of educational knowledge graph from the perspective of "artificial intelligence +": connotation, technical framework and application [J]. Journal of Distance Education, 2019, 37(4): 42-53.
李振, 周东岱, 王勇. “人工智能+”视域下的教育知识图谱: 内涵、技术框架与应用研究[J]. 远程教育杂志, 2019, 37(4): 42-53.
- [28] LIU Y Y, LI Y. Review of medical knowledge map research [J]. Software Guide, 2019, 22(5): 241-247.
刘悦悦, 李燕. 医学知识图谱研究综述[J]. 软件导刊, 2023, 22(5): 241-247.
- [29] DU C. Research and implementation of Leshan tourism intelligent question answering system based on knowledge graph [J]. Computer Knowledge and Technology, 2023, 19(29): 46-48.
杜春. 基于知识图谱的乐山旅游智能问答系统研究及实现[J]. 电脑知识与技术, 2023, 19(29): 46-48.
- [30] ZHAO S, YANG W X, WANG Q Z, et al. A question and answer system based on potato disease and pest knowledge graph [J]. Agricultural Engineering, 2023, 13(8): 29-37.
赵赛, 杨婉霞, 王巧珍, 等. 基于马铃薯病虫害知识图谱的问答系统[J]. 农业工程, 2023, 13(8): 29-37.
- [31] WANG S Y, QIU J T, HONG C Y, et al. Research on online product Q&A based on knowledge graph [J]. Journal of Chinese Information Processing, 2020, 34(11): 104-112.
王思宇, 邱江涛, 洪川洋, 等. 基于知识图谱的在线商品问答研究[J]. 中文信息学报, 2020, 34(11): 104-112.
- [32] WANG L. An improved knowledge graph question answering system for English teaching [J]. Mobile Information Systems, 2022, 2022: 3401074.
- [33] YU S Q, PENG Y, LU Y. An artificial intelligence assistant system for educating people: the structure and function of "AI educator" [J]. Education Research, 2019, 25(1): 25-36.
余胜泉, 彭燕, 卢宇. 基于人工智能的育人助理系统——“AI好老师”的体系结构与功能[J]. 开放教育研究, 2019, 25(1): 25-36.
- [34] SHEN J, QIAO S J, HAN N, et al. Personalized recommendation model with multiple information fusion [J]. Journal of Chongqing University of Technology (Natural Science), 2021, 35(3): 128-138.
沈杰, 乔少杰, 韩楠, 等. 融合多信息的个性化推荐模型[J]. 重庆理工大学学报(自然科学), 2021, 35(3): 128-138.
- [35] WANG D Q, YIN H Y. Research on design of personalized exercise recommendation system based on knowledge graph [J]. China Educational Technology, 2019(17): 81-86.
王冬青, 殷红岩. 基于知识图谱的个性化习题推荐系统设计研究[J]. 中国教育信息化, 2019(17): 81-86.
- [36] WANG L, ZHOU D D, ZHI Y, et al. Research on the construction and application model of the academic teaching system for subject ability development assessment [J]. China Educational Technology, 2019(1): 28-34.
王磊, 周冬冬, 支瑶, 等. 学科能力发展评学教系统的建设与应用模式研究[J]. 中国电化教育, 2019(1): 28-34.
- [37] SHI Y F, PENG H C, TONG M W. Research on generative recommendation strategy of accurate personalized learning path based on learning portrait [J]. China Educational Technology, 2019(5): 84-91.
师亚飞, 彭红超, 童名文. 基于学习画像的精准个性化学习路径生成性推荐策略研究[J]. 中国电化教育, 2019(5): 84-91.
- [38] MOU Z J. Multimodal learning analysis: new growth points in learning analysis research [J]. E-education Research, 2020, 41(5): 27-32, 51.
牟智佳. 多模态学习分析: 学习分析研究新增长点[J]. 电化教育研究, 2020, 41(5): 27-32, 51.
- [39] KE Q C, LIN J, MA X F, et al. The construction direction and development path of digital education resources in the era of new education infrastructure [J]. E-education Research, 2021, 42(11): 48-54.
柯清超, 林健, 马秀芳, 等. 教育基建时代数字教育资源的建设方向与发展路径[J]. 电化教育研究, 2021, 42(11): 48-54.