

基于教育知识图谱的知识建构元空间设计与实现

张义兵¹, 高兴启¹, 满其峰¹, 胡金艳², 许馨予¹

(1.南京师范大学教育科学学院, 江苏 南京 210097;

2.河南师范大学教育学部, 河南 新乡 453007)

[摘要] 跨越社区边界的知识共享是计算机支持下的协作学习的知识创造来源之一,这对重视知识创造过程的知识建构教学来说尤其重要。如何建立跨越边界的知识建构元空间,成为近年来国际跨知识建构社区信息交互研究的重要问题之一;但由于其实现过程受到人工规划、设计与实现的约束,多处于搭建概念模型的研究阶段。本研究基于教育知识图谱的理念,运用 LDA 主题抽取算法抽取知识建构社区的核心主题、发现跨社区的观点之间的联系,进而采用可视化的知识图谱形式构建知识建构元空间,并运用计算机实验加以验证。实验结果表明,教育知识图谱形式下的知识建构元空间在跨社区的观点的核心主题抽取方面以及跨社区观点之间的关系与呈现的功能上表现优秀,主题抽取的准确率为 94.29%,观点与主题之间的关系抽取准确率为 92.85%,同一主题下的观点关系的平均准确率为 89.79%。知识建构元空间的建构尝试,不仅可以为国际知识建构相关研究探索新路径,也可以在实践上更好地支持跨社区的信息交互与共享。

[关键词] 知识建构;教育知识图谱;LDA;跨社区

[中图分类号] G434 **[文献标志码]** A

[作者简介] 张义兵(1967—),男,江苏连云港人。教授,博士,主要从事学习科学与技术、计算机支持下的协作学习(CSCL)、知识建构的研究。E-mail:zhyb304@126.com。

一、引言

随着教育技术以及学习科学的发展,越来越多的计算机支持下的协作学习理论的有效性得到验证,更多的创新教学法成为可能。知识建构(Knowledge Building)教学通过线上与线下相结合的教学环境,培养学生在协作学习环境中进行创新性探究,并在此基础上培养学生知识创造的能力^[1]。这种创新学习的关键在于学生在知识论坛(Knowledge Forum)中创建主题视窗,每个主题下的成员构成了相应的学习社区,社区成员以短文(Note)的形式发表自身观点。在这个过程中,学生在卡尔·波普尔所说的世界3中工作,不断对观点进行改进与升华;由此,社区成员是知识的创造者,而他们的学习成了知识创造过程的副产品^[2]。

在各种计算机支持下的协作学习过程中,跨越学习社区边界的知识共享都是知识创造的重要来源^[3]。因此,跨社区的知识共享对重视知识创造的知识建构教学来说尤为重要。但在知识建构教学中,社区成员在同一个主题下极易形成固定小组,社区交互可能会局限于与小组成员对不同的观点进行改进;随着交互活动的增多,会导致社区成员间的意见逐渐趋同、固化,系统的开放性则逐渐降低。这种现象与 Johan S. G 发现的近些年多学科的大量同质化科研论文会严重阻碍科学研究发展的结论相吻合^[4]。在知识建构教学的研究中,对于如何打破固化僵硬的协作组织的研究已经有了一定的进展:例如,陈羽洁通过研究知识建构社区中组外中间人的特征,提出了引导组外中间人以促进社区深度知识建构的方法^[5],还有学者通过网络

互评教学实践,从评价入手来促进学生的深度学习⁶。但已有研究仍然只关注于促进单个社区内的知识建构协作行为,而对跨社区的协作学习缺乏关注。这种跨社区交互的现实需求和知识论坛相对封闭性的现状之间的矛盾,是阻碍知识建构教学发展的主要技术因素。现有的知识论坛对于跨社区的信息交互缺乏技术支持,致使教学中的跨社区交互行为的促进无从谈起。因此,如何实现便利的跨知识建构社区的信息交互成为近些年知识建构研究的重点,知识建构元空间的概念正是在这种寻求突破的过程中诞生的。

知识建构元空间是一个以跨社区交流为导向的新概念,指向一个对不同社区观点重新组织以促进社区成员跨社区交互的虚拟空间。现有的知识建构元空间实现方法是依靠社区成员制定规则,由学生对知识建构社区中的已有观点进行总结。例如,有学者提出由社区成员按照一定的规则,总结“超级笔记”(Super Notes)构建知识建构元空间的方式以促进跨社区交互⁷。另有学者通过引导两个社区的成员对自己社区观点重新组织再发表在第三个社区空间的方法构建知识建构元空间。这类方法建构的知识建构元空间往往不能对原有社区进行整体性审视,对边缘性观点缺少关注。对于知识建构元空间中不同社区间观点的关系,也采用的是人工分类联系的方式,不能准确地将不同社区的观点联系起来。因此,如何利用教育知识图谱准确抽取跨社区观点的核心主题,以及如何发现与呈现跨社区观点的关系是知识建构元空间构建与发展的关键。

在教育领域中,教育知识图谱由于其对实体与关系抽取的准确性较高,因此,该技术在知识检索、知识问答、学习路径分析、资源推荐等方面得到广泛应用⁸。基于知识建构元空间对跨社区信息交互的需求,以及教育知识图谱在知识抽取与可视化呈现的优势,本研究提出了一种基于教育知识图谱的知识建构元空间实现方案,希望由此可以对单个社区的观点进行全面审视,总结出社区的核心主题;并以核心主题为中心,构建出能够呈现社区间观点相互关系的知识建构元空间。

二、理论基础

(一)知识建构元空间

知识建构元空间的学理基础由来已久,最早可以追溯到知识建构理论萌芽阶段,知识建构理论创始人玛琳·斯卡达玛莉亚和卡尔·布莱特提出运用网络平台促进社区交互,以实现学校向知识建构机构转变的设想⁹。而后,许多学者开始不断完善其构想。例如,张

建伟曾提出以构建观点认知地图(Idea Thread Mapper)的方式促进跨社区成员的观点交互¹⁰,并于2019年进一步提出元空间的概念¹¹。在2021年的知识建构暑期年会(Knowledge Building Summer Institute)中,知识建构元空间的概念引起全世界知识建构研究者的关注。

知识建构元空间的概念关键在于“元”概念的界定。2021年,“元宇宙”(Meta-verse)的概念引发全球性的讨论,其内涵融合人工智能、VR、AR、游戏引擎等虚拟现实技术,展现了在现实世界之上再造世界的可能。由此可见对“元”的概念的探索愈发重要,追溯其源头,“元”概念生自内省法的自我证明悖论的解释,即:客体层面是关于客体本身的表述,而元层面则是关于客体层面表述的表述¹²。在此基础上,我们就可以对知识建构元空间的概念进行解释:首先,从功能上来说,正如元认知是“认知的认知”,其功能在于对自身认知过程的监控与调节,元空间是“空间的空間”。知识建构元空间建立在各个知识建构社区空间之上,通过跨社区的信息整合,促进跨社区的信息交互,进而调整不同社区的认知发展方向,促进观点的改进与社区知识的形成,是社区层面的“元”的实现。其次,从观点改进过程上来说,知识建构元空间是在知识论坛中的社区基础上形成的。学生在原始社区中经过反思、综合、升华等一系列的观点改进工作,不断涌现出新的主题、层次的观点。对于知识建构元空间,原始社区的观点是元空间构成的基础,但元空间不是各个社区堆叠而成,而是每个观点改进方向的总结与归纳,也就是观点层面的“元”。观点层面的“元”是进行社区层面的“元”的基础。因为在知识建构社区中,随着观点的不断改进,社区成员间内部的交互逐渐陷入瓶颈,单个观点改进与涌现的速度变慢¹³,但与此同时,与另外相似主题下的社区的相同点也逐渐增多,这种相同点的增多代表着两个社区观点改进路径的交汇,也代表着知识建构元空间的涌现。

(二)教育知识图谱

知识图谱的理论最早可追溯到1965年,Price指出运用知识图谱分析科学研究引文网络以预测某领域科学研究发展方向的可能性¹⁴。经过后来学者对三维参考模型等知识图谱因素的研究,知识图谱的理念得到了进一步的发展¹⁵。2012年,Google首次正式提出了知识图谱的概念。

知识图谱的实质是结构化的认知语义知识网络。知识图谱中的每个节点代表实体或者概念,节点之间的连线代表实体间的语义关系。知识图谱的最小单位

是一对节点与其中的连线构成的三元组^[16]。知识图谱在科学研究中的应用主要有两种:一种是作为一种科学计量方法,被广泛应用在分析科学研究的前沿方向;另一种是和特定领域结合在一起,形成具有领域特征的领域知识图谱。本研究下的知识图谱就是指在教育领域应用的教育知识图谱。

教育知识图谱在知识的主题抽取与关系发现中早有应用,例如,有学者利用知识图谱的路径分析与关联分析能力,通过对学习者在不同情境下生成的学习路径与推荐路径的对比,构建了基于知识图谱的多约束学习路径推荐算法^[17]。有学者针对现有教育领域模型知识内容分散、能力刻画不足的问题,提出了能够建立知识、问题、能力三者间映射关系,由知识图式、问题图式、能力图式三层图式和知识内容、关联关系、映射关系、学习路径四个要素组成的教育知识图谱 KQA 模型^[18]。有研究者在学习元平台的基础上,采集学生学习过程数据,构建学习认知地图概念模型,在构建学生个性化知识图谱的基础上,叠加学生的认知过程,以期根据学生的认知特征进行个性化的资源和路径推荐^[19]。还有学者从知识建构与联通主义的视角出发,在学习元平台的基础上,设计了以知识网络和社会网络为双核驱动的连接式知识建构模型,构建知识建构活动中的知识与知识、人与知识、人与人之间的可视化连接^[20]。

显然,教育知识图谱在准确抽取跨社区观点的核心主题,以及跨社区观点的关系如何构建的问题上,具有人工实现所不能媲美的优势。一是在观点遍历的全面性上,人工只能对规模较小的观点进行分辨;当观点总数与观点的字数超过了个人的处理极限,教育知识图谱就可以发挥其优势,利用机器学习算法对全部的观点进行遍历,进而解决这个问题。二是在观点关系的准确性上,人工处理会受到个人经验、情感偏向等的价值影响,而教育知识图谱则只会根据关键词频率、主题分布概率等客观信息进行评判。三是在观点组织的科学性上,人工仅能形成一个只能增加节点、不能删改的网络结构,而教育知识图谱则可以根据原始社区观点的实时状况进行修改。因此,教育知识图谱在知识建构元空间的构建上有其独特的优势。

三、知识建构元空间设计

按照一般知识图谱的处理过程,知识建构元空间的实现可以分为三步描述:

首先,是原始数据的获取与预处理。这一步是因为知识建构社区中学生的观点是多元化的,虽然以文

本为主,但也包括音频、视频、图片等多种形式。因此,需要人工对文本以外的观点进行清洗筛选。

其次,是利用教育知识图谱提取社区核心主题与发现观点联系。知识建构元空间的构建是在观点“元”过程的基础上,对社区的“元”的实现,因此,在实体抽取过程中,要进行每个观点的实体抽取,再以此为基础对整个社区进行实体抽取。

常用的知识实体抽取方法包括:基于词典和规则的方法、基于统计机器学习的方法与基于深度学习的算法。对于实体关系挖掘的方法包括有监督的学习、无监督或半监督的学习以及模式匹配方法^[21]。考虑到知识论坛数据源中,学生观点中包含太多生成性领域实体,很难预测一个社区中学生的观点会涉及什么主题,因此,传统情况下对数据进行中文分词、停用词过滤等等操作时候选择的停用词库、中文分词方法等会遗漏一部分频率低的实体,所以本研究采用 LDA 主题模型对观点进行主题提取,并计算语义间的余弦相似度以提高准确性。LDA 是常见的主题模型之一,是一类无监督的学习算法,在训练时不需要手工标注的训练集,仅需要文档集以及指定主题的数量。在算法学习的过程中,参数的初始化相当先验概率,通过样本训练,最终得到真实的后验概率。在观点的关系上,知识建构元空间中的观点可能来自不同社区,因此,本研究中的观点关系不需要分辨观点间的观点类别,只需要判断关系的存在,通过判断每个观点与核心主题的关系,再对不同观点进行余弦相似度计算以准确发现观点间的联系。

最后,是知识图谱的存储与显示。通过上述实体和实体关系挖掘之后,就可以获取知识实体及其关系,本研究运用 echarts 图数据库来存储知识实体及其语义关系。通过 echarts 图数据库构建无指向性的力导向图。echarts 图数据库构建图表类别丰富,可以适用于各种各样的功能。

(一)社区观点数据采集与预处理

既有研究往往采用人工清洗数据的方式,将数据规范化后再用软件进行进一步处理。例如日本 Oshima 团队开发的知识构建话语资源管理器 (Knowledge Building Discourse Explorer, 简称 KBDeX) 根据对话的进度动态绘制每个节点,通过关键词筛选来促进社区间的群体交流,向社区成员推荐下一步的观点改进方向^[22]。其数据源需要人工筛选关键词再进行数据转换后,才能使用 KBDeX 处理。本研究采用爬虫算法对社区中的观点数据进行收集,再将收集后的数据进行人工清洗,按照一定的数据规范,转化为所

需的 json 文件后,再输入知识建构元空间进行知识图谱的提取。

(二)基于 LDA 的主题抽取

隐含狄利克雷分布(Latent Dirichlet Allocation, 以下简称 LDA)由 Blei 于 2003 年首次提出,是一种基于 PLSI (Probabilistic Latent Semantic Indexing)模型延伸而出的主题模型^[23]。LDA 打破了传统的词语—文档的两层结构,形成了一个由文档、主题、词语构成的三层贝叶斯概率图模型。作为一个典型的词袋模型,其主要思想在于:一篇文档由多个独立存在的主题构成,每个主题又由多个词语组成^[24]。在本研究中,学生的每个观点被单独视为一个文档。LDA 主题模型如图 1 所示。

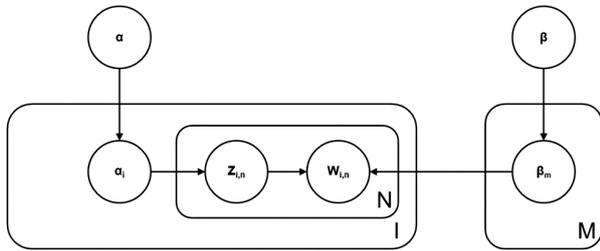


图 1 LDA 主题模型

其中,α、β 是超参数,分别代表 M、N 维向量,其中 α 是每个观点的主题分布的先验分布 Dirichlet 的参数,β 是每个主题下的词语分布的先验分布 Dirichlet 的参数。I 代表观点数量,其实际应用取值为 $i \in [1, I]$,M 代表观点中的主题数量,其实际应用取值为 $m \in [1, M]$,N 代表词语数量,其实际应用取值为 $n \in [1, N]$ 。α_i 是一个 I*M 的矩阵,代表第 i 个观点中的 m 个主题的概率分布矩阵。β_m 是一个 M*N 的矩阵,代表第 m 个主题中的 n 个词语的概率分布矩阵。Z_{i,n} 为第 i 个观点中的第 n 个词语的主题编号,W_{i,n} 为第 i 个观点中的第 n 个词的概率分布。

针对以上的 LDA 模型,主题词抽取过程大致如下:

(1)对每个观点计算主题分布概率:

$$\vec{\alpha}_i \sim \text{Dir}(\vec{\alpha}) \quad \text{公式(3.1)}$$

(2)对每个主题进行词语采样:

$$\vec{\beta}_i = \text{Dir}(\vec{\beta}) \quad \text{公式(3.2)}$$

(3)对每个抽取单词 W_{i,n} 从主题分布 α_i 中选择主

题编号:

$$Z_{i,n} \sim \text{multi}(\vec{\alpha}_i) \quad \text{公式(3.3)}$$

(4)抽取核心词语的多项分布:

$$W_{i,n} \sim \text{multi}(\vec{\beta}_{z_{i,n}}) \quad \text{公式(3.4)}$$

依此上,可得总体的主题分布为:

$$P(\vec{Z}_{i,n} | \vec{\alpha}) = \prod_{i=1}^I P(\vec{Z}_{i,n} | \vec{\alpha}) = \prod_{i=1}^I \frac{\Delta(\vec{n}_i + \vec{\alpha})}{\Delta(\vec{\alpha})} \quad \text{公式(3.5)}$$

其中 n_i 表示第 i 个观点中各主题中词的个数的集

合。

总体的词语分布为:

$$P(\vec{W}_{i,n}, \vec{Z}_{i,n} | \vec{\alpha}, \vec{\beta}) = P(\vec{W}_{i,n} | \vec{Z}_{i,n}, \vec{\beta}) = P(\vec{Z}_{i,n} | \vec{\alpha}) = \prod_{m=1}^M \frac{\Delta(\vec{n}_k + \vec{\beta})}{\Delta(\vec{\beta})} \prod_{i=1}^I \frac{\Delta(\vec{n}_i + \vec{\alpha})}{\Delta(\vec{\alpha})} \quad \text{公式(3.6)}$$

其中 n_k 表示第 k 主题中词的个数的集合。

在获取总体词语分布后,对于如何确定最优主题,学界目前倾向采用两种方式,一种是根据主题一致性 (Topic Coherence) 来判断效果的优劣,另一种是 Blei 提出的基于困惑度的方法。本研究采用主题一致性的方式来判断效果的优劣,越靠近 1 则证明效果越好。

(三)教育知识图谱的呈现

教育知识图谱的重点在于其节点以及节点之间关系的构建与呈现,这个过程通过分别对社区以及社区中的每个观点进行基于 LDA 的主题词抽取实现。教育知识图谱的实体发现过程是对每个观点进行主题抽取以确定其观点的核心关键词。社区的核心主题词作为每个实体簇的核心节点,每个观点为簇的围绕节点,将实体与关系通过 echarts 库构建为力导向图形式的知识图谱。其中,每个实体簇的核心并不是观点内容的原本重复,而是基于社区观点内容提取出来的,代表了社区最根本的特征和社区观点改进方向的总结。

四、知识建构元空间的构建实验及效果分析

(一)实验数据来源

本研究选取了中文版知识论坛平台中 A、B 两个社区的观点作为数据来源,观点总数 900 余条,总字数超过 50 万。在实验之前对实验采用的两个社区进行了考察,确定了两者之间存在内容联系的可能。

(二)实验过程

1. 数据预处理

首先针对社区中的观点内容,爬取其标题、内容,

21	topic 1	topic 2	topic 3	69	topic 1	topic 2	topic 3	121	topic 1	topic 2	topic 3		
22	0	学习	软件	进行	70	0	概念	学生	122	0	评语	鼓励	学科
23	1	设计	研究	方法	71	1	自我	效能	123	1	表现形式	开头	实践
24	2	建构	知识	数据分析,	72	2	建构	学习	124	2	方法	评价	常见,
25	topic 1	topic 2	topic 3	73	topic 1	topic 2	topic 3	125	topic 1	topic 2	topic 3		
26	0	知识	建构	问题	74	0	维度	具体	126	0	评价	家长	教师
27	1	发展	世纪	多样化	75	1	自我	效能	127	1	评价	学生	之间
28	2	真实	认知	社区,	76	2	效能	自我	128	2	学生	自评	互评,
29	topic 1	topic 2	topic 3	77	topic 1	topic 2	topic 3	129	topic 1	topic 2	topic 3		
30	0	问题	配送	存储	78	0	个体	因素	130	0	地区	我国	气温
31	1	外卖	公司	了解	79	1	三元	班杜拉	131	1	影响	气候	问题
32	2	学生	货源	培养,	80	2	自我	效能	132	2	划分	人口迁移	命名,

图2 观点主题抽取部分结果

转化为 json 串格式。再遍历 json 串中标题和内容,进行分词和清洗。对标题等确定的内容直接进行分词处理,对观点具体内容进行逐句分词。最后将分词结果进行清洗,运用中文停用词词表等停用词库删除停用词与空值。进入下一步的 json 串应包括“topic”“content”“category”“topic_seg”“content_seg”五个属性,分别代表了观点的标题、内容、社区编号、标题分词结果、内容分词结果。

2. 主题抽取

由于计算机无法直接识别处理文本内容,且 LDA 是基于统计学实现的词袋模型,因此,需要将文本内容转换为一个二维的词频矩阵。研究采用 DF 技术构建词频矩阵以反映词语的重要程度,并在此基础上将文本转换为向量格式。而简单的分词方式,不能区分文本中词语的重要程度,因此引入 TF-IDF 对观点进行分词的基础上进行 LDA 主题抽取。

主题抽取结果分为两部分,分别是可能包含的主题及其概率分布情况。首先对每个观点进行主题抽取,得到每个观点的主题矩阵以及相匹配的概率矩阵。这个过程的处理,可以有效地对观点进行识别,确定每个观点的主要内容,为进一步对观点进行分类做准备。如图 2 所示,根据抽取结果,人工大致浏览是否存在明显数据异常,如有则需要对 LDA 主题抽取过程的参数进行调整。

在每个观点的主题抽取结果基础上,再以相似过程对整个社区进行主题抽取。在每个观点的基础上,对整个社区进行整理,最终得到整个社区中的主题词和每个观点与对应每个主题的概率分布情况。基于观点的概率分布结果,对观点进行归类,首先基于概率分布的大小,进行初步判断,取其中的最大值作为该观点归属的指标,判断其归属主题。再对相同主题下的各观点进行余弦相似度的计算,以各观点之间的距离判断其相关性,舍弃同一主题下,相关性较弱的观点,对观点的主题归属进行进一步修正。

3. 可视化呈现

可视化呈现共分三步。首先构造节点数据,代表主题与观点,因此节点分为两类,一类是主题节点,一

类是观点节点。然后构造边的数据,代表主题与观点之间的联系。结合 echarts 库,将节点与边的数据进行可视化呈现,绘制教育知识图谱,如图 3 所示。

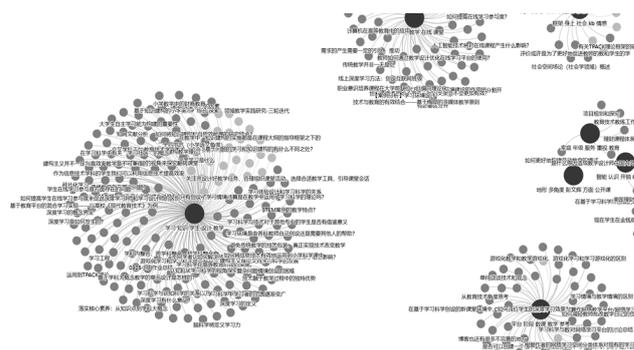


图3 知识图谱形式下的知识建构元空间

图中不同主题分别成簇,其中大节点代表核心主题,深灰色小节点代表其中一个社区的观点,浅灰色小节点代表另一个社区的观点。其中核心主题节点上悬浮显示该主题包含的主题词。观点节点上悬浮显示该观点的标题,图谱可以通过缩放对节点进行更细致的显示,每个观点可以显示其标题以方便定位,点击观点可以查看该观点的具体内容,如图 4 所示。



图4 知识图谱下观点的具体内容

(三) 实验效果分析

针对本研究提出的两个关键问题,对实验结果形成的知识建构元空间进行三方面的验证:知识图谱中的核心主题抽取是否正确,观点与主题之间的关系是否准确,同一主题下的观点是否存在联系。

因为观点内容的复杂及关系的复杂性,难以计算其实体关系的准确率与召回率,只能人工判断观点与主题之间的关系是否准确。本实验中共有观点 904 条,最后形成主题簇共 35 个。

本文定义 P_1 作为主题抽取的准确率, T_c 表示正确抽取的核心主题, T 表示核心主题总数。该准确性描述的是抽取出来的主题词是否为有意义的知识实体,是否符合该社区的讨论热点。其计算公式如下:

$$P_1 = \frac{T_c}{T} \quad \text{公式(4.1)}$$

定义 P_2 作为观点与主题之间的关系是否准确的准确率, I_c 表示与核心主题关系正确的观点数, I 表示观点总数,其计算公式如下:

$$P_2 = \frac{I_c}{I} \quad \text{公式(4.2)}$$

定义 P_3 作为同一主题下的观点是否存在联系的平均准确率, D_c 表示与同一主题下的存在关系的观点对, D 表示同一主题下的观点对的总数,其计算公式如下:

$$P_3 = \frac{\sum_{i=1}^T \frac{D_c}{D_i}}{T} \quad \text{公式(4.3)}$$

经过人工校对与计算,主题抽取的准确率 P_1 为 94.29%,观点与主题之间的关系的准确率 P_2 为 92.85%,同一主题下的观点联系的平均准确率 P_3 为 89.79%。知识建构社区本身所具有的数据稀疏性依然限制了核心主题的提取以及观点之间关系的归类。例如,社区发表观点“我支持你的想法”“我觉得你的实验不错”“我不同意你的观点”等无领域实体的观点,也限制了主题抽取的准确性。观点稀疏性减少了频繁项集的数量,进一步限制了核心主题的出现和观点联系的构建。但本实验证明了,基于教育知识图谱的知

识建构元空间在跨社区观点的核心主题抽取方面以及跨社区观点之间的关系与呈现的功能上表现优秀,能够更加准确地支持跨社区的信息交互与共享。

五、结 语

从理论上来说,知识建构理论重视跨社区的信息交互与共同创造;但是,这种创造不是某个社区成员对其他社区观点的简单的、累加式的阅读,而是对不同社区间有关系的观点建立相互联系后所形成的新观点的涌现。在知识建构教学实践中,不同学段、不同学科的跨社区的知识建构历程也证明了跨社区的信息交流能够有效促进社区间的观点改进过程。但是,仅有理论模型或者仅仅靠人工,难以处理知识建构跨社区中数据量极大的观点和复杂的关系。本研究提出的知识建构元空间,就是运用教育知识图谱对跨社区的内部关系进行重新组织后的可视化呈现。

本研究利用 LDA 主题提取模型抽取实体与实体关系,构建教育知识图谱以实现知识建构元空间,实验结果显示优秀。在国际知识建构研究中,本研究是试图将设想转化为可操作的方法的一种尝试,可以说是一个从理论到技术实现的前导性研究。对于知识建构教学实践而言,本研究的方法可以在一定程度上作为跨社区知识建构实践的支持技术来使用。当然,研究仍有很大的进步空间,因为观点提取中的数据稀疏性依然对主题抽取的过程有一定影响,且 LDA 主题抽取模型本身在短文本的处理上具有一定局限性。现有的知识建构元空间仅仅是观点组织网络的可视化功能的实现,进一步的知识建构元空间应该可以实现知识建构支撑环境的其他功能,例如,观点的发表、评论、引用以及升华。如何在现有的基于教育知识图谱的知识建构元空间基础上实现观点的改进是进一步的研究方向。

[参考文献]

- [1] SCARDAMALIA M, BEREITER C. Knowledge building: theory, pedagogy, and technology [M]. New York: Cambridge University Press, 2006.
- [2] 张义兵, 陈伯栋, Scardamalia M, et al. 从浅层建构走向深层建构——知识建构理论的发展及其在中国的应用分析[J]. 电化教育研究, 2012, 33(9): 5-12.
- [3] 张义兵, 满其峰. 知识建构共同体中两种协作脚本的组间交互差异研究[J]. 电化教育研究, 2015, 36(8): 5-10, 17.
- [4] JOHAN S G C, JAMES A E. Slowed canonical progress in large fields of science[J]. Proceedings of the national academy of sciences, Oct 2021, 118(41): 1-5.
- [5] 陈羽洁, 张义兵, 徐朝军. 知识建构社区外组中间人的形成特征及作用研究[J]. 电化教育研究, 2020, 41(2): 38-44.
- [6] 张义兵, 孙俊梅, 木塔里甫. 基于知识建构的同伴互评教学实践研究[J]. 电化教育研究, 2018, 39(7): 108-113.
- [7] YUAN G J, ZHANG J. Connecting knowledge spaces: enabling cross-community knowledge building through boundary objects[J].

- British journal of educational technology, 2019,50(5): 2144–2161.
- [8] 张翔翔, 马华. 知识图谱与图嵌入在个性化教育中的应用综述[J]. 计算机系统应用, 2022, 31(3): 48–55.
- [9] SCARDAMALIA M, BERETER C. Computer support for knowledge–building communities[J]. Journal of the learning sciences, 1994, 3(3): 265–283.
- [10] ZHANG J W, TAO D, CHEN M H, et al. Co–organizing the collective journey of inquiry with idea thread mapper[J]. Journal of the learning science, 2018, 27(3): 1–63.
- [11] ZHANG J W, YUAN G J, BOGOUSLAVSKY M. Give student ideas a larger stage: support cross–community interaction for knowledge building[J]. International journal of computer–supported collaborative learning, 2020, 15(4): 389–410.
- [12] 汪玲, 方平, 郭德俊. 元认知的性质、结构与评定方法[J]. 心理学动态, 1999(1): 6–11.
- [13] 蒋纪平, 胡金艳, 张义兵. 知识建构学习社区中“观点改进”的发展轨迹研究[J]. 电化教育研究, 2019, 40(2): 21–29.
- [14] PRICE D J. Networks of scientific papers[J]. Science (New York, N.Y.), 1965, 149(3683): 510–515.
- [15] KRETSCHMER H. Coauthorship networks of invisible colleges and institutionalized communities [J]. Scientometrics, 1994, 30(1): 363–369.
- [16] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582–600.
- [17] ZHU H P, TIAN F, WU K, et al. A multi–constraint learning path recommendation algorithm based on knowledge map [J]. Knowledge based systems, 2018, 143: 102–114.
- [18] 钟卓, 唐焯伟, 钟绍春, 赵一婷. 人工智能支持下教育知识图谱模型构建研究[J]. 电化教育研究, 2020, 41(4): 62–70.
- [19] 万海鹏, 余胜泉. 基于学习元平台的学习认知地图构建[J]. 电化教育研究, 2017, 38(9): 83–88, 107.
- [20] 万海鹏, 余胜泉, 王琦. 连接式建构: 知识建构研究的新取向[J]. 电化教育研究, 2021, 42(10): 12–18, 24.
- [21] 薛增灿. 知识论坛的教育知识图谱构建与可视化系统实现[D]. 南京: 南京师范大学, 2021.
- [22] MATSUZAWA Y, OSHIMA J, OSHIMA R, et al. Kbdex: a platform for exploring discourse in collaborative learning [J]. Procedia – social and behavioral sciences, 2011, 26(1): 198–207.
- [23] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. Journal of machine Learning research, 2001, 3(4/5): 993–1022.
- [24] 刘三女牙, 彭晔, 刘智, 等. 面向 MOOC 课程评论的学习者话题挖掘研究[J]. 电化教育研究, 2017, 38(10): 30–36.

Design and Implementation of Knowledge Building Meta–space Based on Educational Knowledge Graph

ZHANG Yibing¹, GAO Xingqi¹, MAN Qifeng¹, HU Jinyan², XU Xinyu¹

(1.School of Education Science, Nanjing Normal University, Nanjing Jiangsu 210097;

2.Faculty of Education, Henan Normal University, Xinxiang Henan 453007)

[Abstract] Knowledge sharing across community boundaries is one of the sources of knowledge creation in computer–supported collaborative learning, which is especially important for knowledge construction teaching and learning that values the process of knowledge creation. How to establish a knowledge building meta–space across boundaries has become one of the important issues in international research on information interaction across knowledge building communities in recent years. However, due to the constraints of manual planning, design and implementation, it is mostly at the research stage of building conceptual models. Based on the concept of educational knowledge graph, this study uses the LDA topic extraction algorithm to extract the core topics of knowledge construction communities, finds the connections among the views across communities, and then constructs a knowledge construction meta–space in the form of a visual knowledge map, and validates it through computer experiments. The experimental results show that the knowledge construction meta–space in the form of educational knowledge graph performs well in extracting the core topics of cross–community views and in the function

of the relationship and presentation between cross-community views. The accuracy of topic extraction is 94.29%, and the accuracy of the relationship between views and topics is 92.85%. The average accuracy of viewpoint relationships under the same topic is 89.79%. The attempt to construct knowledge construction meta-space can not only explore new paths for international knowledge construction-related research, but also better support cross-community information interaction and sharing in practice.

[Keywords] Knowledge Building; Educational Knowledge Graph; Latent Dirichlet Allocation (LDA); Cross-community

(上接第 68 页)

strategies, peer emotional challenge strategies and their interaction significantly affected the quality of online collaborative interaction content, network structure and overall assessment scores. Moreover, peer emotional interaction strategies significantly affect learning satisfaction. This study shows that learners perform best in online collaborative interactions under the interaction between the peer argument atmosphere and the peer emotional evaluation strategy. Therefore, in the process of online collaborative interaction with challenging situations, the use of peer emotional interaction strategies, especially peer emotional evaluation strategies, can better guide learners to continuously and deeply engage in the construction of collaborative knowledge.

[Keywords] Peer Emotional Interaction Strategy; Online Collaborative Learning; Social Sharing of Emotion; Social Network Analysis; Content Analysis

(上接第 76 页)

[Abstract] Based on theories of interactive ritual chain, pedagogy, social psychology, etc., this study takes the bullet screen text of the online open course "*Pedagogy*" loaded on the Bilibili website as the research object, combines qualitative and quantitative research methods to crawl and analyze the bullet screen of the website using natural language processing technology. It was found that in the process of online teaching, "the learning subject has a strong ability to recreate; the teaching part has sufficient motivation to explore; the frequency of interaction in the teaching process is high; and the positive evaluation of teaching feedback is high". However, there are also challenges such as "the deviation from the predetermined goal which needs to be faced, the stability of the willingness to explore which needs to be maintained, the diversity of the participating subjects which needs to be consolidated, and the advanced nature of the feedback objects which needs to be empowered". It is suggested that the quality of online teaching should be improved through strategies such as "guidance + facilitation, high-level + challenge, preset + company, temperature + brightness" to respond to questions such as "what is the quality and effect of online teaching?".

[Keywords] Bullet Screen Text; Online Teaching; Effective Strategy