

DOI:10.13718/j.cnki.xsxb.2022.09.017

数字人文视域下民国乡村师范教育文献的 语义组织与智慧应用研究^①

章雷¹, 胡蓉², 唐振贵³, 贺彩云⁴, 王琼宇⁴

1. 西南大学 图书馆, 重庆 400715; 2. 西南大学 教师教育学院, 重庆 400715;
3. 广西财经学院 新闻与文化传播学院, 南宁 530007; 4. 西南大学 计算机与信息科学学院, 重庆 400715

摘要: 民国乡村师范教育实践不仅能为当下乡村教师教育的内涵式发展提供历史依据, 也能为中国特色乡村教师教育的本土化创生带来现实启示. 重庆作为民国乡村教育实践汇聚之地, 珍藏有该段历史实践的丰富史料. 本研究借助西南大学图书馆馆藏优势, 从数字人文视角切入, 以构建民国乡村师范教育本体 RTEduRP 为基础, 针对民国乡村师范教育文献进行语义组织, 搭建民国乡村师范教育文献知识库 RTEduRPDocKB 平台, 探索基于语义检索和语义增强技术的智慧应用. 研究成果能为挖掘民国乡村师范教育的当代价值提供数据基础设施和便捷利用机制, 进而实现民国乡村师范教育文献的“活化”应用.

关键词: 民国乡村师范教育文献; 数字人文; 本体; 知识库; 语义组织; 智慧应用

中图分类号: G659.29

文献标志码: A

文章编号: 1000-5471(2022)09-0116-09

On Semantic Organization and Smart Application of Rural Teacher Education Documents in Republican Period from Perspective of Digital Humanities

ZHANG Lei¹, HU Rong², TANG Zhengui³,
HE Caiyun⁴, WANG Qiongyu⁴

1. Southwest University Library, Southwest University, Chongqing 400715, China;
2. College of Teacher Education, Southwest University, Chongqing 400715, China;
3. School of Journalism and Cultural Communication, Guangxi University of Finance and Economics, Nanning 530007, China;
4. College of Computer and Information Science, Chongqing 400715, China

Abstract: The practice of rural teacher education in the Republican Period(RTEduRP)can not only provide historical basis for the connotative development of current rural teacher education, but also bring practical enlightenment to the localization of rural teacher education with Chinese characteristics. As the gathering

① 收稿日期: 2022-07-28

基金项目: 重庆市基础研究与前沿探索项目(cstc2018jcyjAX0149), 中央高校基本科研业务费专项资金项目(XDJK2019C118), 西南大学创新研究 2035 先导计划项目(SWUPilotPlan027).

作者简介: 章雷, 硕士, 主要从事科学计量分析、民国文献数字化等方面研究.

place of RTEduRP, Chongqing treasures rich historical materials of this historical practice. Taking advantage of the collections of the Southwest University Library, this study starts from the perspective of digital humanities based on the construction of the ontology of RTEduRP, conducts semantic organization for the documents of RTEduRP, builds the RTEduRPDocKB platform, and explores smart applications based on semantic retrieval and semantic enhancement techniques. The research results can provide data infrastructure and convenient utilization mechanism for mining the contemporary value of RTEduRP, and then realize the “activation” application of RTEduRP.

Key words: rural teacher education documents in the Republican period; digital humanities; ontology; knowledge base; semantic organization; smart application

2021年2月,中共中央办公厅、国务院办公厅印发了《关于加快推进乡村人才振兴的意见》,明确强调要加强乡村教师队伍建设。乡村教育是乡村振兴战略的基础性工作,教育兴则乡村兴。乡村教育振兴关键在人,在于乡村教师教育对于“合格”乃至“卓越”乡村教师的培养。而如何推动乡村教师教育事业全面发展的话题可以追溯到20世纪20~30年代兴起并兴盛的民国乡村师范教育,期间以陶行知、黄质夫等教育家为代表的乡村师范教育实践,不仅为当下乡村教师教育的内涵式发展提供了历史依据,更为中国特色乡村教师教育的本土化创生带来现实启示。

对于民国乡村师范教育的整体性考察得益于对相关文献的系统把握,更依赖于对那段尘封历史的深入挖掘。重庆作为民国乡村教育实践汇聚之地,珍藏有该段历史实践的丰富史料,其中不乏海内外珍本。作为一种技术逻辑与人文逻辑相耦合的全新研究范式,数字人文(Digital Humanities)为民国乡村师范教育文献内容的语义组织与智慧应用带来契机,通过深入文献内容层面的语义组织,可为挖掘民国乡村师范教育的当代价值提供数据基础设施和便捷利用机制。为此,本研究以构建民国乡村师范教育本体为基础,针对民国乡村师范教育文献进行语义组织,探索其智慧应用场景,实现民国乡村师范教育文献的“活化”应用。研究成果将实现一源多用,可为国家或地区乡村教师教育政策的制定提供便捷化借鉴与参考,服务乡村振兴;为相关学术研究提供平台与数据支撑,服务乡村教育研究;也可为公众提供智慧化的历史记忆与文化共享传播服务,服务历史文化传播。

1 民国乡村师范教育及其相关文献概述

1.1 民国乡村师范教育

民国乡村师范教育是一个历史概念,萌芽于1912年中华民国临时政府成立时对文化教育的全面改革。随着五四运动时期乡村教育思潮的兴起,国内教育家和教育团体逐渐意识到创立乡村师范教育是解决乡村问题的良方^[1]。1919年余家菊首先倡导乡村教育运动,其方向就是“向师范学校去运动”,具体而言就是要在师范学校各科教授中着眼于乡村,设置乡村教育学科,创立乡村实验学校,并养成师范生服务乡村社会的精神^[2]。1921年袁希涛、黄炎培等人成立了义务教育期成会,认为中国的广大民众在乡村,推行义务教育必须着眼于乡村,开展乡村教育,为此,必须大力培养乡村师资。在该会所办《义务教育》刊物上,人们纷纷撰文要求设立乡村师范,探讨乡村师范教育的样式,如《办理农村师范学校之管见》(9期)、《筹设乡村师范学校之意见》(15期)、《办理农村师范学校的商榷》(20期)、《办理农村师范的旨趣》(24期)等^[3]。此后乡村师范教育经历了从理论准备走向实践探索的过程。典型的乡村师范教育实践如黄质夫从1923年起先后创办的江苏界首乡村师范学校、江苏栖霞乡村师范学校、浙江湘湖乡村师范学校、贵州省立贵阳乡村师范学校(后改名为国立贵州师范学校);晏阳初1924年以保定地区作为实验区进行乡村教育实验,对乡村教师实施培训;而陶行知1927年领导创办的晓庄师范学校(以下简称晓庄学校)则将乡村师范教育推向高潮。不过,直到1928年大学院召开第一次全国教育会议提出《整顿师范教育制度案》,乡村师范学校才被明确列入

师范教育制度中^[3], 获得其在师范教育制度中的合法地位. 总体上, 20 世纪 20~30 年代是中国乡村师范教育兴起与发展的重要时期, 本文所指民国乡村师范教育也主要涉及此时期.

崔运武在《中国师范教育史》中将乡村师范教育界定为“是以一种办在乡村, 以培养乡村小学教员为基本目标的师范教育, 就其教育程度而言, 是一种中等师范教育”^[3]. 教育家古棣在《乡村师范概要》中明确表示“设在都市的师范学校不能造就适合乡村发展需要的师资, 乡村师范学校宜与都市师范学校分离, 乡村师范教育要培养能发展乡村教育之人”^[4]. 为此, 本文所探讨的民国乡村师范教育主要指 20 世纪 20~30 年代办在乡村, 以培养合格乡村小学师资和改造乡村社会为目的的中等师范教育.

1.2 民国乡村师范教育文献

《辞海》对“文献”一词的解释有三: 一指典籍与贤者; 二指具有历史价值的图书文物资料, 亦指与某一学科有关的重要图书资料; 三指如今记录有知识的一切载体的统称, 即用文字、图像、符号、声频、视频等手段以记录人类知识的各种载体. 本研究所指民国乡村师范教育文献, 主要是指记录了民国时期(尤其是 20 世纪 20~30 年代乡村师范教育发展兴盛时期)乡村师范教育相关人物、理论、实践等, 具有突出历史价值与史料价值的图书、报刊、文集等资料. 上述资料在本研究团队所在西南大学图书馆馆藏较为丰富, 团队前期已整理出版的《民国乡村教育文献丛刊三编》(全三十册)和《民国乡村建设文献丛刊》(全六十五册), 以及馆藏的《民国乡村教育文献丛刊》(全二十八册)、《民国乡村教育文献丛刊续编》(全三十四册), 为研究提供了有力的文献保障. 本研究拟通过整合数字人文的技术逻辑与人文逻辑, 深入挖掘这一时期涉及乡村师范教育的文献中所承载的乡村师范教育管理、理论与实践等内容, 以语义组织方式提供便捷与智慧化的文献知识服务, 为新时代乡村教师教育提供借鉴与参考.

2 数字人文在教育相关领域的探索与实践

2.1 数字人文的研究范式

数字人文(Digital Humanities)源于人文计算(Humanities Computing)^[5], 致力于围绕人文社会科学领域研究对象本体, 实现与之相关的各类数字资源的深度整合与保存, 向用户提供专题信息服务并为相关应用提供支持^[6], 其一端是高效的计算, 另一端是人文沟通^[7]. 数据密集、工具支持、跨界合作是数字人文研究最鲜明的三大特征. 作为一种技术逻辑与人文逻辑相耦合的新兴跨学科研究热点领域, 数字人文正从广度和深度两个维度重构人文社会科学研究^[8]. 数字人文的技术体系涉及数字化技术、数据内容加工技术、数据分析技术、可视化技术、VR/AR 技术与机器学习等技术: ①数字化技术如扫描、拍摄、OCR 识别等; ②数据内容加工技术如本体建模、文本图像编码与语义描述、命名实体提取等, 其中本体的建立支持概念关系的管理, 支持语料库的智能扩展, 支持文本自动分析与语义分析, 因而成为数字人文建设中非常重要的知识组织工具; ③数据分析技术如文本分析、内容挖掘、时序分析、地理空间分析、社会关系分析等; ④VR/AR 技术如人机交互技术、互动测量、游戏化学习等, 在数字人文成果的展示与传播方面发挥着重要作用; ⑤机器学习技术如自动分类、图像视频音频识别和分析、个性化服务、精准推送等^[9]. 整个技术体系由底层到高级, 体现出数字人文的实现逻辑, 其中, 本体的构建是数字人文项目的知识组织基础, 在此基础上构建的知识库则将形成数字人文项目的信息资源基础, 基于本体和知识库则可实现各种智慧应用开发.

2.2 教育相关领域的数字人文研究现状

近年来, 世界各地陆续成立数字人文研究机构并开展数字人文项目, 全球范围内有近 200 个以“数字人文”命名的中心、项目、实验室、团队或圈子. 相关数字人文项目如美国布朗大学数字学术中心的“加里波第与意大利统一运动档案”^[10], 纽约公共图书馆数字画廊的“地图整经机”^[11], 美国弗吉尼亚大学历史研究中心完成的“影谷项目”^[12], 国家图书馆“华夏记忆项目”^[13], 中国科学院计算机研究所、武汉大学与浙江大学联合开展的“数字敦煌”项目^[14], 武汉大学信息资源研究中心董慧教授主持开发的“中华基本史籍分析系

统”, Coursera 推出的北京大学“中国古代文化”, 上海交通大学的“中医药与中华传统文化”, 上海图书馆开发的“家谱知识库系统”“盛宣怀档案库”“上海年华”, 中华书局开发的“中华经典古籍库”^[15], 浙江大学图书馆开发的“民国文献大全”^[16]等。

如前所述, 在数字人文项目中, 本体构建是知识组织基础. 具体到教育领域, 国内外相关研究也在探索各种教育领域本体的构建. 例如, 文献[17]构建了教育领域顶层本体, 将该本体与相关教育网站和教育信息系统联合, 可以实现异构信息资源的语义检索; 文献[18]以教育语义网为背景, 提出基于本体的自适应学习系统架构, 为学习者提供基于个人知识空间和偏好的学习服务; 文献[19]构建了教育技术学专业图书领域本体, 为专业课程的发展提供智能服务; 文献[20]构建了基于关联课程数据与知识点本体的知识图谱, 增强了教学资源间的语义关系; 文献[21]构建了涵盖学校、学生、试卷、实体、知识点及各种测评指标等元素的教育测评知识图谱, 包括了基于本体技术的模式层构建和依托于模式层结构定义的数据层构建; 文献[22]构建了课程及教学大纲本体, 并将其应用于课程和教学大纲相关知识的语义搜索, 实现智能服务; 文献[23]构建了面向场景的 MOOC 资源本体, 并探讨了该本体模型在语义检索、知识导航和个性化场景推荐方面的应用. 国外的教育领域本体研究主要涉及了课程建模与管理本体、描述学习领域的本体、描述学习者数据的本体、描述 e-learning 服务的本体、描述学习领域与学习者数据的多本体^[24].

综上, 尽管目前已有多样化的教育本体, 但均不适合描述民国乡村师范教育文献. 首先, 教育领域的通用型顶层本体不适合描述具体专题或是微观层面的教育现象; 而当下多数教育专题本体主要基于现代教育体系来构建, 例如教学本体、课程本体等, 这些本体提供的类及实例不能体现出民国乡村师范教育的特点. 其次, 上述基于现代教育体系来构建的本体, 缺乏从历史的视角进行描述, 即已有本体主要是教育本体, 不是教育史本体, 而从本质上看, 民国乡村师范教育本体应彰显教育史研究的特点. 第三, 现有教育本体主要从教育知识体系的角度构建, 忽略了知识使用者的研究与利用需求和作为知识生产来源的教育史料之间的密切关系, 未能在二者之间搭建起供需匹配的桥梁. 鉴于此, 本研究拟从数字人文视角切入, 结合民国乡村师范教育及其文献特征, 以及教育史研究与利用需求, 在构建民国乡村师范教育本体基础上, 探索对民国乡村师范教育文献的语义组织与智慧应用.

3 民国乡村师范教育本体的构建

3.1 本体构建原则与方法

通过调研民国乡村师范教育文献和教育史研究者的需求, 本研究发现:

1) 民国乡村师范教育在教育管理(如教育政策制定、教育经费投入、教育推广方式)、教学内容(如师范类和农业类相关课程的设置)、教学方法(如教学做合一的方法)、实习要求(如涵盖教育实习和农事实习)等方面有着突出的特点, 成为构建民国乡村师范教育专题本体的重要元素.

2) 教育史研究中强调对原始资料的引证溯源, 因此需要在民国乡村师范教育本体中添加有关溯源的类, 需在本体中考虑例如民国时间、民国时期教育的分期、教育历史地理等方面.

3) 当前民国乡村师范教育本体的构建并非要寻求一种大而全的宏观思路, 而是可以先从学者们对民国乡村师范教育史研究的主题需求中选定和凝练语义化的主题, 如涉及教育管理、教学、课程、实习等方面; 同时, 本体的构建要从史料文献的形式特点出发来思考能够为这些研究主题提供什么样的信息, 进而从知识服务的角度来处理知识供求的问题. 因此, 本研究中本体构建的总体原则是兼顾教育史研究需求与民国乡村师范教育文献的内容和形式特征, 构建一个初步满足学者研究需要的专题本体, 而本体内容的丰富与完善则可通过后期迭代来实现.

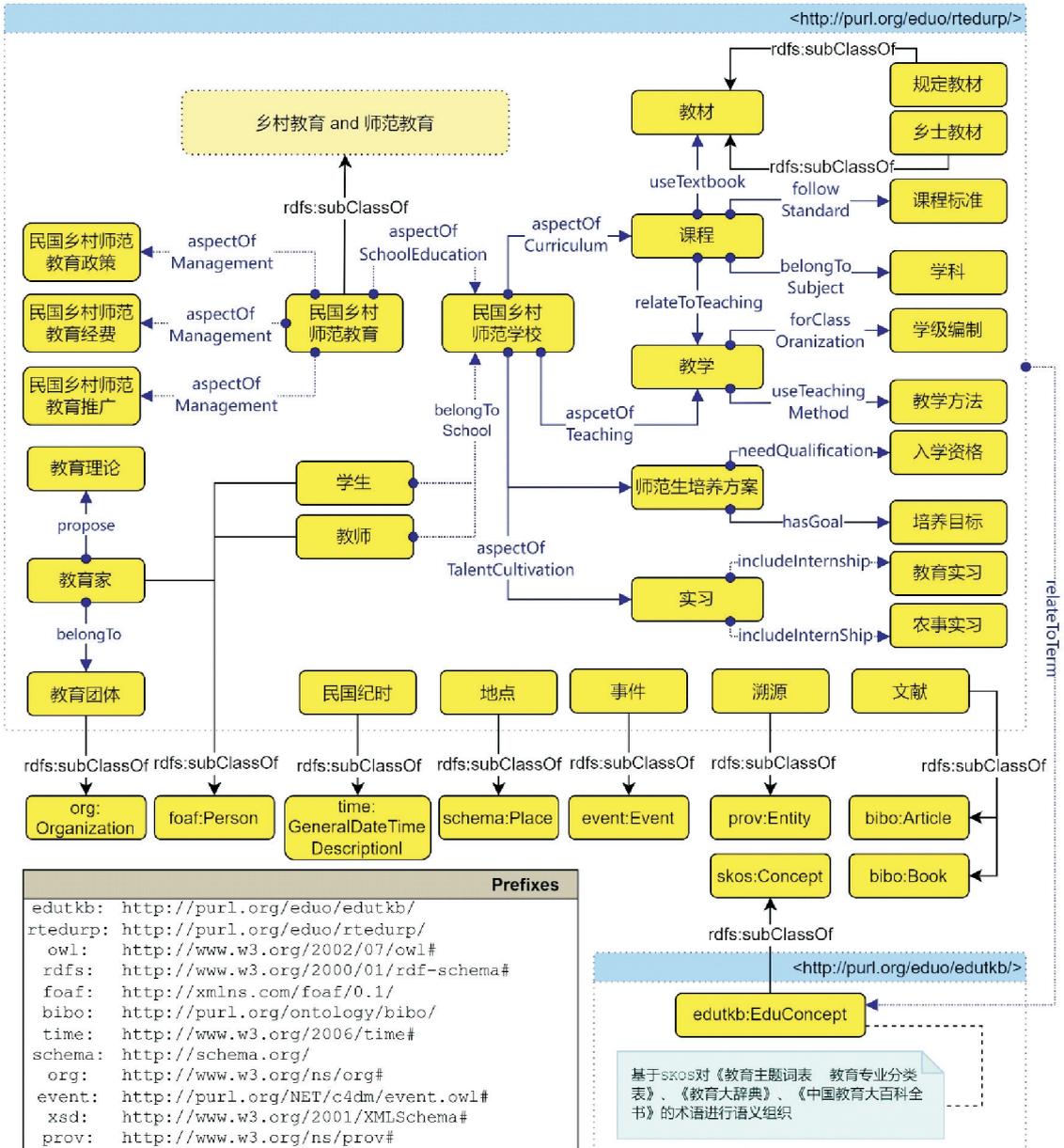
本体是一个开放集成的概念体系, 是共享概念模型的形式化规范说明, 能够在语义和知识层次上描述信息. 领域本体构建方法各不相同, 国内外较为成熟的本体构建方法有 IDEF-5 法、TOVE 企业建模法、Methodology 法、循环获取法、七步法和骨架法等. 鉴于七步法具有最强的本体建模表达能力^[25], 并被广

泛应用于本体建模中,具有较强的成熟性和普适性,本研究选择七步法构建本体.具体而言,在明确构建民国乡村师范教育领域本体基础上,复用现有本体、借助教育主题词表构建概念术语、定义类之间的层次关系和属性间关系,并定义属性约束,创建实例.

3.2 民国乡村师范教育本体的描述

如图 1 所示,民国乡村师范教育(Rural Teacher Education in the Republican Period)本体(简称 RTEduRP 本体)兼顾宏观、中观与微观三层视角,采用 Graffoo^[26] 本体图形框架来表示.

从宏观视角看,RTEDuRP 本体以民国教育研究的教育管理、课程与教学等主题为中心,通过名如“aspectOf*”格式的对象属性来关联各类主题实体,如 aspectOfCurriculum, aspectOfManagement 等.每一主题通过分类层次关系进一步细分,也为中观视角和微观视角的知识分类组织,以及文献的主题标引和知识导航提供语义支撑.



图例: 类 (黄色方框) 类约束 (虚线方框) 谓词 (声明类公理) (实线带箭头) 对象属性 (实线带圆点) 便捷对象属性 (虚线带圆点)

图 1 民国乡村师范教育本体 RTEduRP

从中观视角看,RTEDuRP本体详细刻画了教育教学中的核心概念及概念间的关系,如学校、课程、教学与教材等实体的特征及其之间的关系。民国乡村师范学校在人才培养方面颇具特色,师范生的培养方案中对入学资格和培养目标通常都有明确规定,例如,晓庄学校以培养乡村儿童和人民所敬爱的导师为培养目标,具体分目标为健康的体魄、农人的身手、科学的头脑、艺术的兴趣、社会改造的精神,因此,该本体中入学资格和培养目标成为师范生培养方案实体的两个重要的对象属性值。在学校课程设置方面,民国乡村师范学校的课程涉及丰富的科目,例如公民、体育、家事、卫生、国文、算学、地理、历史、生物、化学、物理、伦理学、劳作(工艺)、美术、音乐、农业及实习、农村经济及合作、水利概要、教育概论、教育心理、小学教材及教学法、小学行政、教育测验及统计、乡村教育、实习等,且各科都有相应的课程标准^[27]。系统性的课程设置将为当下乡村教师教育相关课程设置带来重要启示,因此,RTEDuRP本体将学科与课程标准作为课程实体的两个重要的对象属性值。此外,鉴于民国时期乡村师范学校与普通师范学校相比有其特殊性,前者重在培养学生毕业后深入农村,办理适合农村需要的教育事业,因此其教育范围以及教育实习范围更广,除教育实习(如参观、见习、试教)外,还涉及农事实习(如劳作教育实习)^[28]。因此,教育实习和农事实习将作为RTEDuRP本体中实习实体的两个对象属性值。在教学方面,民国乡村师范学校基于农村教育较为通行的学级编制(如单级编制)和教学方法(如教学做合一)实施师范生教学,例如,特别强调培养师范生对于单级小学的行政设施、编制、管理、教材、教法等问题的研究,以及体现教学做合一的生活法。由此,针对的学级编制和采用的教学方法将作为教学实体的重要对象属性值。在课程使用教材方面,针对民国乡村师范教育实际情况,区分出规定教材和乡土教材两个子类。

从微观视角看,RTEDuRP本体主要描述民国乡村师范教育中的人物、时间、地点、事件等基本实体,以增强或扩展民国乡村师范教育文献在文本内容方面的语义解释。具体而言,本研究基于W3C的Time本体构建顺序坐标结构的民国纪年表示形式,通过继承time:GeneralDateTimeDescriptionl建立民国纪时的时间描述,同时建立起民国时间与公历时间的对应;此外,采用民国分期表示从不同教育研究维度、以时间段表示的民国时期的教育发展阶段划分。地点实体涉及民国地名、行政区划的属性(如归属关系、层级、别名、方位等),同时建立与现代行政区划的关联。人物实体在foaf:Person基础上增补人物的曾用名、字、号等别名;针对民国时期的教育家类,建立其教育理论、教育经历、教育活动等方面的描述。事件实体对事件进行命名与描述,并与相关的人物、时间和地点进行关联。除此之外,教育团体、教育文章和著作复用已有本体的org:Oranization, bibo:Article和bibo:Book等类;溯源实体用于描述实体的定义或实体的属性值所依据的史料文献来源,体现出RTEDuRP本体的溯源功能。

在关联方面,本研究通过继承、映射(owl:equivalentClass, owl:equivalentProperty, owl:sameAs)和特定关联属性等多种方法,将RTEDuRP本体与教育术语知识库、外部知识库进行关联。其中教育术语知识库(EduTKB, Educational Term Knowledge Base)是基于SKOS本体来组织教育类的分类主题词表、教育辞典和教育大百科中的术语词汇,通过描述这些术语的概念和结构构建术语知识库。

最后形成的当前版本的RTEDuRP本体共包含145个类、154个数据属性、167个对象属性。

4 民国乡村师范教育文献的智慧应用案例

4.1 民国乡村师范教育文献的语义组织

民国乡村师范教育文献的语义组织成果体现为相关知识库的构建,构建流程如图2。

首先,构建民国乡村师范教育知识库RTEDuRPKB。使用Topbraid Composer来构建RTEDuRP本体,包括类、属性及SHACL(shapes constraint language)约束;选择民国教育相关的学术研究文献《中国行政区划通史·中华民国卷》^[29]、教育大辞典中的词条(如人物、事件等文本)等外部知识来源,通过机器半自动抽取和专家构建的方法来构建RTEDuRP本体的实例数据;并将部分人物、书目等数据与上海图书馆的数字人文开放数据平台进行关联。

其次,构建教育术语知识库 EduTKB. 将词典通过抽取、转换后存入关系数据库,再使用 python 编写转换程序,用 SKOS 来组织术语的名称、定义,表示术语上位、下位、相关等语义关系以及相关匹配、精确匹配等映射属性.

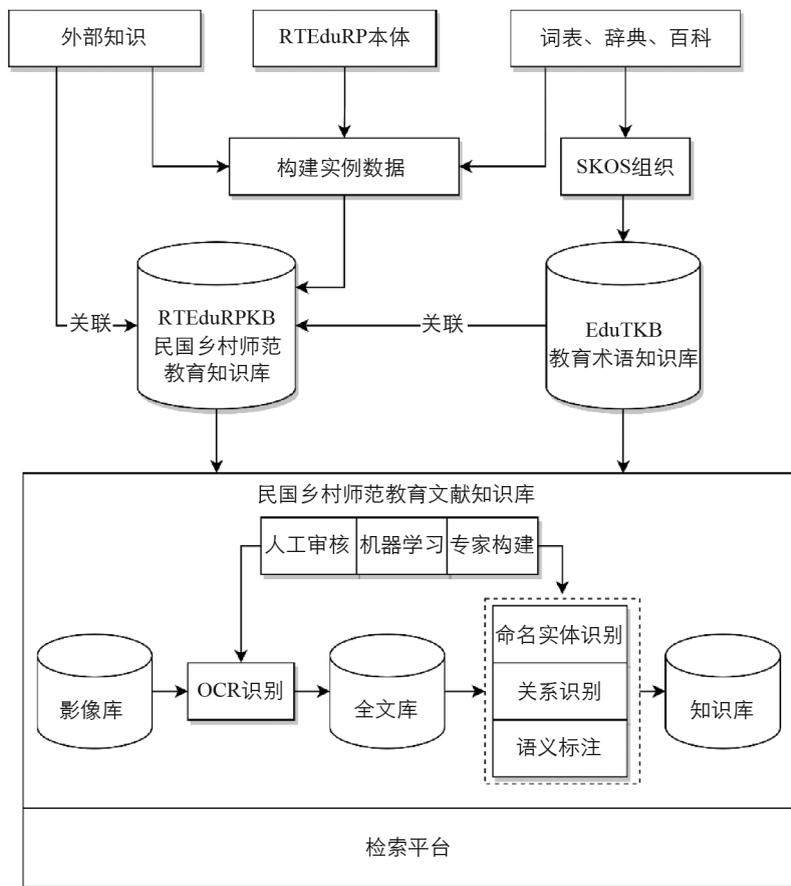


图 2 民国乡村师范教育文献的语义组织

第三,构建民国乡村师范教育文献知识库 RTEDuRPDocKB. RTEDuRPDocKB 包括原始文献的影像库、文本全文库和语义知识库,目前收录范围涉及前文所述馆藏的民国乡村教育文献丛刊及其续编和三编. 具体而言,通过 OCR 识别和人工校对的方法将影像转换成文本格式的全文库;通过机器自动识别和专家构建两种方式从文本中进行命名实体识别、关系识别和语义标注,构建基于 RTEDuRP 本体描述的语义知识库,同时建立该知识库与 RTEDuRPKB 和 EduTKB 两个知识库的关联. 其中,机器自动识别采用基于 huggingface 的预训练模型的 finetune 方法实现,主要对人名、地名、时间和事件进行初步识别,再通过人工审核来确认识别的结果. 对于教育领域专门的实体,如课程、教学方法等实体及其关系,本研究搭建专门的语义标注平台,由专家构建结合专业知识和本体进行实体及其关系的语义标注.

最后,构建检索平台. 将知识库的 RDF 数据存储在 neo4j 图数据库中,全文文本和 RDF 数据使用 solr 构建索引. 前端检索界面采用 vue 开发, RDF 可视化功能采用 LodLive 组件实现,后台应用服务采用 node.js 技术实现.

4.2 民国乡村师范教育文献的语义检索与语义增强

语义增强(Semantic Enrichment)是一种通过应用语义技术来增强数据价值的策略,可应用于各行各业各种类型的数据,以提高数据质量、可发现性和再使用性,在智慧数据建设中具有重要意义^[30]. 文献[31]将语义增强的方法与技术归为语义搜索和浏览、语义中介、语义标注、语义分析和知识发现 4 类. 本研究在前述对民国乡村师范教育文献进行语义组织基础上,搭建民国乡村师范教育文献知识库(RTEDuRP-

DocKB)平台,提供简单检索、全文检索、语义检索等多种检索方式.语义检索的结果记录提供影像视图、全文视图和知识视图 3 种查阅视图,其中:影像视图提供文献原件的影印图像;全文视图提供文本形式的内容,并对文本中特定概念实体标注其语义信息,通过关联知识库的相关知识,实现对文本的语义增强;知识视图以 RDF 可视化的方式呈现实体(结点)及其关联(关系),用户可以根据结点上的聚焦、展开、关闭等操作图标来进行知识导航.上述 3 种视图可以按平铺模式排列,通过联动设置,同步呈现影像、文本、知识的互动.图 3 所示的案例展示了“农事课程”语义检索结果的其中一条记录,通过三视图同时呈现出该实体所在的史料原文、文本上下文以及知识关联.其中,全文视图呈现出相关扩展信息,知识视图中则呈现出该课程与学校、地点、时任校长等关联信息.总体上看,平台的语义检索和语义增强功能初步实现了民国乡村师范教育文献的一种智慧应用.



图 3 民国乡村师范教育文献知识库中的语义检索与语义增强示例

5 结语

教育是国之大计、党之大计,教师是立教之本、兴教之源,乡村教师教育是乡村振兴战略的基础性工作.从数字人文视角,对民国乡村师范教育文献展开语义组织与智慧应用研究,将实现该类文献的“活化”应用,为挖掘民国乡村师范教育的当代价值提供数据基础设施和便捷利用机制.基于此驱动力,本研究在构建 RTEduRP 本体,进而构建相关知识库基础上,搭建了 RTEduRPDocKB 平台,通过语义检索和语义增强功能,初步实现了民国乡村师范教育文献的智慧应用探索.未来的研究可着眼于以下几方面:

1) 丰富与完善 RTEduRP 本体.当前本体主要聚焦于教育管理、课程、教学等主题,并且对教育管理方面仅作了宏观主题描述.未来可进一步拓展相关主题,并着力从中观和微观角度对本体加以细化.

2) 扩大知识库规模.鉴于民国乡村师范教育主题研究的专业性,目前研究中自动化技术的使用较为有限,人工审核和专家构建工作较为费时费力,增加了研究成本,也导致数据加工规模不大.未来在探索提升自动化技术应用程度的同时,可考虑通过众包方式构建大规模文献知识库.

3) 构建更加丰富的应用体系.本研究所构建的本体与知识库当前主要应用于语义检索,未来可考虑探索构建一整套数字人文方法(如利用社会网络、大数据和可视化方法),实现对民国乡村师范教育文献的系统、深入挖掘.

参考文献:

- [1] 谭玉秀, 战萌. 20 世纪二三十年代中国乡村师范教育探研 [J]. 吉林师范大学学报(人文社会科学版), 2017, 45(1): 88-92.
- [2] 余家菊. 乡村教育运动的涵义和方向 [J]. 中华教育界, 1921, 10(10).
- [3] 崔运武. 中国师范教育史: 中国教育史专题研究丛书 [M]. 太原: 山西教育出版社, 2006.
- [4] 古棣. 乡村师范概要 [M]. 上海: 商务印书馆, 1936: 39.
- [5] WILLARD M C. Humanities Computing [M]. New York: Palgrave Macmillan, 2005.
- [6] 赵生辉, 朱学芳. 数字人文仓储的构建与实现 [J]. 情报资料工作, 2015(4): 42-47.
- [7] UNSWORTH J. What is Humanities Computing, and What is Not? [J]. Jahrbuch für Computerphilologie, 2002(4): 71.
- [8] 刘炜, 叶鹰. 数字人文的技术体系与理论结构探讨 [J]. 中国图书馆学报, 2017, 43(5): 32-41.
- [9] 刘炜, 谢蓉等. 面向人文研究的国家数据基础设施建设 [J]. 中国图书馆学报, 2016, 42(5): 29-39.
- [10] Resources in the Garibaldi Panorama [EB/OL]. (2013-05-06)[2022-02-15]. <https://library.brown.edu/cds/garibaldi/resources/>.
- [11] JOSH J. New York Public Library Puts 20, 000 Hi-Res Maps Online & Makes Them Free to Download and Use [EB/OL]. (2018-02-03)[2014-03-31]. <http://www.openculture.com/2014/03/new-york-public-library-puts-20000-hi-res-maps-online.html>.
- [12] The Valley of the Shadow [EB/OL]. [2022-02-15]. <http://valley.lib.virginia.edu/VoS/choosepart.html>.
- [13] 华夏记忆 [EB/OL]. [2022-02-15]. <http://www.nlc.cn/old2010hxjy/syjt/>.
- [14] 数字敦煌 [EB/OL]. [2022-02-15]. <https://www.e-dunhuang.com/index.htm>.
- [15] 中华经典古籍库 [EB/OL]. [2022-02-15]. <http://www.gujilianhe.com.cn/>.
- [16] 民国文献大全(~1949) [EB/OL]. [2022-02-15]. <http://cadal.hytung.cn/>.
- [17] 刘小乐. 教育领域顶层本体构建及其应用研究 [D]. 长春: 吉林大学, 2012.
- [18] 欧阳杨. 教育语义网中基于本体的自适应学习系统建模 [D]. 杭州: 浙江大学, 2008.
- [19] 夏波. 教育技术专业图书领域本体构建研究 [D]. 曲阜: 曲阜师范大学, 2012.
- [20] 张萌. 课程知识图谱组织与搜索技术研究 [D]. 武汉: 武汉大学, 2016.
- [21] 罗明. 教育测评知识图谱的构建与应用 [D]. 北京: 北京工业大学, 2019.
- [22] 彭潇. 课程及教学大纲的本体构建及其在语义搜索的应用 [D]. 杭州: 浙江工商大学, 2019.
- [23] 时家昊. 面向场景的 MOOC 资源本体构建及应用研究 [D]. 保定: 河北大学, 2020.
- [24] STANCIN K, POSCIC P, JAKSIC D. Ontologies in Education-State of the Art [J]. Education and Information Technologies, 2020, 25(6): 5301-5320.
- [25] 杨秋芬, 陈跃新. Ontology 方法学综述 [J]. 计算机应用研究, 2002, 19(4): 5-7.
- [26] FALCO R, GANGEMI A, PERONI S, et al. Modelling OWL Ontologies with Graffoo [M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2014: 320-325.
- [27] 杜成宪. 民国乡村教育史料丛刊-第六册 [M]. [影印本]. 北京: 国家图书馆出版社, 2014.
- [28] 杜成宪. 民国乡村教育史料丛刊-第二十八册 [M]. [影印本]. 北京: 国家图书馆出版社, 2014.
- [29] 傅林祥, 郑宝恒. 中国行政区划通史·中华民国卷 [M]. 上海: 复旦大学出版社, 2007.
- [30] 曾蕾, 谭旭. 数据的语义增强——解读图档博支持数字人文的新动向 [J]. 数字人文研究, 2021, 1(1): 65-86.
- [31] DAMJANOVIC V, KURZ T, WESTENTHALER R, et al. Semantic Enhancement: the Key to Massive and Heterogeneous Data Pools [C]//Proceeding of the 20th International IEEE ERK Conference. New York: IEEE Press, 2011: 413-416.