

关联数据学习平台建设研究^{*}

范炜 张丽

(四川大学公共管理学院, 成都 610064)

摘要: 关联数据学习平台的建设有助于普及关联数据教育, 提升关联数据学习效能, 促进关联数据资源的开发与利用。以开发与利用为导向, 针对关联数据知识体系庞大、学习资源相对分散、技术门槛相对较高、理论学习与技术实践脱节等学习问题, 构造突破单一线性学习思路的探究式关联数据学习场景, 分析场景构成要素和学习资源基础。从关联数据的能力指标体系设计、学习资源描述与组织以及支撑探究式学习功能等方面深入剖析关联数据学习的当前最佳实践LD4PE。关联数据学习平台建设是一个复杂且持续的系统工程。以LD4PE为平台建设基础, 将LD4PE的自建学习路线、自选学习资源组合、辅助实践操作、鼓励学习分享互动等作为主要抓手, 适时推进对中文环境下关联数据学习平台的本地化建设。

关键词: 关联数据; 探究式学习场景; 学习平台; LD4PE

中图分类号: G255.76

DOI: 10.3772/j.issn.1673-2286.2019.12.006

在开放数据生态中, 关联数据 (Linked Data) 已经成为数字资源表征与发布的一种重要形式。众多信息资源拥有机构使用关联数据对外发布自有资源, 并鼓励第三方应用和最终用户使用这些开放数据资源。

在世界范围内, 关联数据理念已经得到广泛普及, 通过关联数据发布的数据集数量不断增长。关联开放数据云图 (Linked Open Data Cloud) 的数据集统计显示, 2007年5月—2019年3月, 其收录的数据集从最初的12个激增到1 239个 (包含16 147个链接), 涉及的主题领域及内容包括政府、生命科学、语言学、媒体、出版、社交网络、用户生成内容以及跨领域内容^[1]。得益于优质内容与高度结构化特点, 随着关联数据集种类与数量的增长, 其资源价值逐渐凸显, 具有深层次的数据分析、挖掘与知识发现价值。学术资源数据集的开放与关联, 让科学家有机会纵深和跨界发现潜在知识关联, 激发研究创新; 数字人文依赖的数据基础很多以关联数据形式发布; 政府发布社会、经济、民生相关的开放数据集, 促进政府信息公开和进一步开放利用; 文化机构通过关联数据发布特色馆藏资源, 以提高其在网络上的可见性和可寻性, 允许多元化分享与混搭创新, 助力

公共文化服务推广, 增强民众文化自信等。

然而, 关联数据的开发与利用面临诸多现实挑战。其中, 许多从事信息资源建设、数字图书馆开发的专业人士和信息管理专业学生缺乏专业化、系统化的关联数据处理能力。关联数据技能学习是开发与利用的前提。

1 关联数据学习场景分析

1.1 学习者及学习困境分析

关联数据的学习者类型较为复杂。关联开放数据云图里各行业主题的信息资源机构从业人员、关联数据应用服务提供商, 以及信息资源管理、信息系统开发、数据分析与挖掘方向的师生都是明确或潜在的关联数据学习者。技术应用的最佳实践是不断进行探索的经验总结。关联数据的实施者“在做中学、学中做”, 他们的技术掌握与应用障碍在一定程度上代表学习者的学习困境。

联机计算机图书馆中心 (Online Computer Library

^{*}本研究得到国家社会科学基金项目“情景应对视角下的情报资源动态响应机制研究” (编号: 14CTQ014) 资助。

Center, OCLC)曾在2014—2018年先后3次开展国际关联数据推广与应用调查^[2]。该调查搜集世界范围内关联数据应用者的进展详情,着重分析被调查机构发布与利用关联数据时出现的问题,并提出建设性解决办法。该调查体现了良好的国际代表性,调查的样本数目在逐次上升,2018年参与国家数达到23个,被调查机构数达到81家,在图书馆、博物馆、研究机构、政府的基础上新增服务提供商,涉及各种类型的信息资源与技术服务机构。一方面,这体现出关联数据的应用范围与推广力度在不断扩大;另一方面,服务提供商开始为客户提供专门的关联数据服务,关联数据已成为一种成熟的行业级技术服务。

该调查给出了关联数据实施者所遇到的共性障碍^[3],实施者在关联数据的理论运用、数据质量、实践工具和技术实现等方面存在理论学习与实践应用中的断层问题,其中技术学习难度高的问题反响一直最为强烈。关联数据学习困境主要表现在3个方面:①需要了解和掌握的知识点相对庞杂,而与之相关的学习资源相对分散,缺乏系统化完整的学习资源;②关联数据涉及一些技术与编程,这些会提高部分非技术类学习者的使用门槛;③关联数据与最终用户之间的交互缺少友好的使用界面(中介工具)。关联数据的查询、分析与挖掘需要使用SPARQL构造查询语句,主要采用编程方法和API调用方式,目前尚未出现更加友好、易用的交互式图形界面。

1.2 新环境下学习理念与学习方式变化

网络时代的内容呈现广泛化、碎片化、开放化和快速迭代等特点,新的学习理念与学习方式不断涌现。联通主义学习理论是现有各种学习理论中贴近新时代环境下学习发生过程与方式观念的,MOOC是当前的最佳实践^[4]。此理论的核心是连接与网络,可解释关联数据学习者以自我的知识网络为导向,在学习关联数据的理论与技术知识的再创造过程中,连接组建新的学习网络。非正式学习指不以明确的组织形式开展的、自发进行的知识与技能学习的过程^[5]。这种自发的主动学习没有时空限制,学习内容具有场景性,学习效果以自身满意度为依据。关联数据学习者在应用场景中触发学习,通过社交活动和网络搜索等多种方式学习关联数据知识,从而达成个人学习目标。

探究式学习最大的特点是“做中学”,即“学以致

用、用以助学”,其核心在于培养学习者主动认知与学习的热情,培养自主推理和求解的能力,培养跨学科知识关联与迁移应用的能力。区别于其他学习,探究式学习强调探究的程序,除了学习知识与提问,更要鼓励学习者开展探究活动,即搜集解释问题的证据并依据事实形成回答,同时学习者之间需要交流论证自己的回答,在不断比较其他可能的、具有科学性的解释中评价并修改自己的回答^[6]。这反映在关联数据这类工程性技术领域的学习场景特征上,即让学习者自主学习关联数据技术知识,共享已掌握的知识技能并表达出来,实施具体发布与消费关联数据的项目(技术实践),并能总结实践经验指导新的实践(技术创新)与分享经验(社区交流)^[7]。根据沈志宏等^[8]提出的关联数据技术体系,学习者的学习任务涵盖关联数据的内容组织、创建发布、浏览检索与互联维护4个方面,学习者除了要进行大量技术实践以掌握相应的技能,还要具备在实际应用中做出最佳决策的能力。

1.3 探究式学习场景要素

关联数据探究式学习场景的发展愿景是建设成为关联数据的探索博物馆,成为关联数据项目实施者的创新基地。学习场景以虚拟的非正式学习环境为基础,把学习者放在教学的中心,同时发挥教育者的指导参与作用。他们可以指导学习者确定学习主题,培养其提问、实验、检验测试与解释结果的探究学习能力,观察学习者的思考状态,并给予帮助。参考目标导向和网络学习环境的设计理念^[9],结合关联数据学习特点,提出探究式关联数据学习场景的5个要素。

(1) 学习目标。学习者背景多元化,对关联数据实践的理论与技术学习目标不一,需求层次不同。明确学习目标才能有针对性地组合学习资源,高效提升关联数据能力水平。

(2) 知识框架。关联数据的知识框架是引导自主学习与探究实践的重要基础,学习者在此基础上可根据自身需求组建个性化知识网络。

(3) 学习资源。指关联数据学习的各种资料,如书籍、技术教程、标准文档、博客文章、在线课程、视频讲座等。

(4) 学习活动。完成特定学习目标的一系列操作,由教育者指定或学习者自行设计。关联数据学习活动主要包括理论学习、应用案例分析和技术操作等。

(5) 虚拟学习社区。由关联数据学习者、教育者共同组成, 分享学习成果与技术创新。

关联数据探究式学习场景的落实体现为学习平台的建设、学习者的自主学习能力和教育者辅助课程设计等。其中, 学习平台的建设是探究式学习场景的具体化承载, 能够提供关联数据知识框架、汇聚开放的关联数据学习资源、可操作的实验环境与社区交流渠道。

1.4 学习资源基础

国外在关联数据教育培训方面已经做了许多探索。网络上分散着各类关联数据学习资源, 包括在线教程类、书籍类、课程类、实践类等。在线教程类资源一般体量较小、篇幅较短, 适合入门者快速了解关联数据或专门学习其中某项技术主题。例如, 剑桥语义公司 (Cambridge Semantics) 开发的在线教程 *Semantic University*, 涵盖语义网技术概述与关联数据技术主题^[10]。书籍类资源包括以关联数据直接命名的书籍和以语义网、RDF、SPARQL 等命名的相关书籍, 主要读者是技术开发人员, 例如 *Evolving the Web into a Global Data Space*^[11]、*Learning SPARQL*^[12]、*Linked Data: Structured data on the Web*^[13] 等。课程类资源直接面向关联数据学习者的系统化学习需要, 提供课程大纲、知识框架以及配套的学习资源, 注重学习过程与反馈。包括由大学主持开发的专门课程或相关课程, 如芬兰阿尔托大学语义计算研究组开发的在线学习平台关联数据学院^[14]和由专门职业培训机构为特定专业人群开发的课程, 如图书馆果汁学院 (Library Juice Academy) 为图书馆员量身定制的付费课程《关联数据导论》, 是 XML 与 RDF 认证资格的 6 门课程之一^[15]。实践类资源指各类信息资源机构开展关联数据实践的项目成果与经验。例如, 万维网联盟 (W3C) 图书馆关联数据孵化组最终报告汇编了各类关联数据应用实践^[16], OCLC 公开了关联数据在图书馆领域的项目成果与相关资料^[17] 等。

除了以上提到的, 欧盟的 EUCLID 项目是一个有效集成各类学习资源的交互式学习平台, 其主要特点有: ①提供模块化关联数据技术知识体系, 并为主流学习者类型提供不同的模块组合指导; ②提供类型丰富的线上学习资源, 这些资源以项目组开发出的 EUCLID 课程教育电子书为核心, 扩展关联数据主题相关的附加资源, 如演讲、出版物、单元拓展阅读的材料、推荐的软件工具和相关培训项目的链接等; ③关注学习效果的课程

设计, 每个模块的学习都设有测试或实践练习; ④通过举行网络研讨会与线下培训活动不断深入、拓展现有学习资源, 提供更多学习机会^[18]。

在众多关联数据学习项目中, 关联数据专业教育 (Linked Data For Professional Education, LD4PE) 是当前最具特色的一个教学研究项目, 也是一个集成各类学习资源的交互式学习平台^[19]。与 EUCLID 不同, LD4PE 的重点是围绕能力组织学习资源并支持探究式学习。如何组合复用已有的众多关联数据学习资源, 满足关联数据学习场景中自主化、个性化、实践性的探究式学习要求, 这是关联数据学习平台建设研究的主要目标。以下对 LD4PE 进行详细剖析, 将其作为当前关联数据学习平台建设的主要基础。

2 关联数据学习平台 LD4PE 分析

2.1 LD4PE 概述

LD4PE 是国际上专门研究关联数据“教与学”的重要项目, 由美国博物馆与图书馆服务机构 (Institute of Museums and Library Services) 资助, 主持机构为华盛顿大学信息学院和都柏林核心元数据倡议 (Dublin Core Metadata Initiative, DCMI), 主要面向美术馆、图书馆、档案馆和博物馆 (Galleries, Libraries, Archives and Museums, GLAM) 领域专业人士的关联数据学习需要。学习关联数据 (Learning Linked Data) 项目是 LD4PE 的先导调查, 所产出的调查报告界定了 LD4PE 所服务的学习群体与其学习需求, 规划出可供探究式学习的关联数据学习平台建设方案, 从目标、设计要点与理念为 LD4PE 奠定了实施基础^[20]。

LD4PE 的目标是开发一个支持关联数据探究式学习, 以能力指标为核心, 组织学习资源的关联数据探索平台 (以下简称“平台”)。LD4PE 重点要解决关联数据学习资源分散、知识系统化欠缺等问题, 满足 GLAM 领域的关联数据学习与培训要求。学习者个体的学习能力有差异, 学习目标侧重点也有差异。例如, 从业人员注重快速学习并指导解决实际问题, 学生注重理论与技术的掌握, 教师则侧重引导指导学生利用该学习平台。

LD4PE 的核心是关联数据能力指标体系 (Linked Data Competency Index, LDCI), 基于 LDCI 组织学习资源, 以关联数据形式发布, 开发个性化学习路线定制工具包, 为用户提供获得社区支持的渠道、构建探究式

关联数据学习平台。

LD4PE设计理念体现为可重用、可扩展与社区驱动。可重用旨在实现教育方式的推广,平台所提供的工具包即用即取,关联数据学习资源的组织方式可重现于其他学科领域。可扩展即保证项目的生命力能够跟上关联数据技术组的继续发展,提供简便工具修改能力指标与学习资源。社区驱动包含两层含义,一层指学习者通过社区进行学习交流,另一层指开源众包的项目工作模式。表1体现了专家与社区参与相结合的工作方式。一是,由专家组工作搭建LDCI与学习资源的基础组织框架;二是,由专家组领导社区力量进行开发协作,负责版本维护和管控。两级工作模式让参与开发LDCI的学习者和专家之间的互动形成对等,达成高度合作,从而实现LDCI可持续发展。

表1 LD4PE工作流程

工作流程项	说明
发现和描述	专家组先尽可能地发现和描述现有的学习资源和关联数据能力
分析讨论	专家组对无法映射到现有能力清单上的学习资源问题以及学习资源与能力指标粒度不统一等问题进行讨论
编辑组会议	编辑组审核能力内容,即界定内容范围,与处理指标能力粒度
征集意见	公开征集关联数据社区成员意见,获取对LDCI不同主题簇下能力的修改意见,利用GitHub进行版本控制,通过编辑组审核,不断改进LDCI

2.2 关联数据能力指标体系LDCI设计

LDCI是LD4PE的核心,也体现出关联数据知识体系。LDCI描述了一组关联数据的学习目标和学习结果,对关联数据从设计到建模的过程中所需的相关知识、技能、实践和思维方式等能力要求进行定义。

LD4PE本身就是一个关联数据应用的良好示范,

项目采用美国国家科学数字图书馆的成就标准网络(Achievement Standards Network, ASN)来开发LDCI。ASN是一个机器可读的美国成就标准存储库,涵盖众多教育机构发布的教学成就标准,其目标在于通过关联数据的发布方式集成美国教育资源与教学知识体系^[21]。成就标准即学科的能力体系,或称知识框架,由一组能力的陈述组成。LDCI使用RDF描述关联数据能力,将能力作为学习成果的陈述,具体采用成就标准描述语言。将LDCI文档与每一个能力陈述当作实体,并定义URL来唯一识别实体,采用成就标准网络、都柏林核心元数据、教育元数据、简单知识组织体系、词汇描述注解等表达出能力的层级结构与语义关系。通过主题簇、主题、能力和基准的陈述之间的关系,陈述与整个能力体系文档的关系,形成对能力的分类单元路径。图1展示了LDCI中一条能力的分类单元路径示意,通过描述主题簇、主题、能力之间的关系,形成从叶遍历到根的树结构。此外,还使用RDF描述了LDCI与ASN中其他知识框架之间的相关性,以及LDCI与学习资源的对应关系,这为学习者创建个性化学习路线打下基础。

LDCI整体为4级展开结构:主题簇、主题、能力、能力要求的基准。其中主题簇是主题的分组;主题是能力的分组;能力代表关联数据学习的需要掌握知识、技能、思维方式等,是要学习的理论知识;基准是证明具有特定能力的技术实践。LDCI在内容上呈现为层级树状,但主题之间没有人为限定的逻辑顺序与重要程度差异。

如图2所示,目前LDCI包括6个主题簇、30个主题和95个能力。主题簇包括RDF基础、关联数据基础、RDF词汇与应用纲要、RDF数据的生成与转换、与RDF数据的交互以及关联数据应用的开发。LDCI涵盖关联数据基础知识和技能要求,但不是完整的是第6部分还没有充分展开,关联数据的存储、分析与利用尚有

<http://achievementstandards.org/resources/D2695955>
关联数据能力指标体系

<http://asn.desire2learn.com/resources/S2696001>
主题: RDF中的识别



<http://asn.desire2learn.com/resources/S2695991>
主题簇: 资源描述框架的基础知识

<http://asn.desire2learn.com/resources/S2731548>
能力: 知道在RDF中所有事物(thing)通过唯一资源识别符URI进行命名,如代理、地点、事件、物件与概念等

图1 LDCI分类单元路径示例

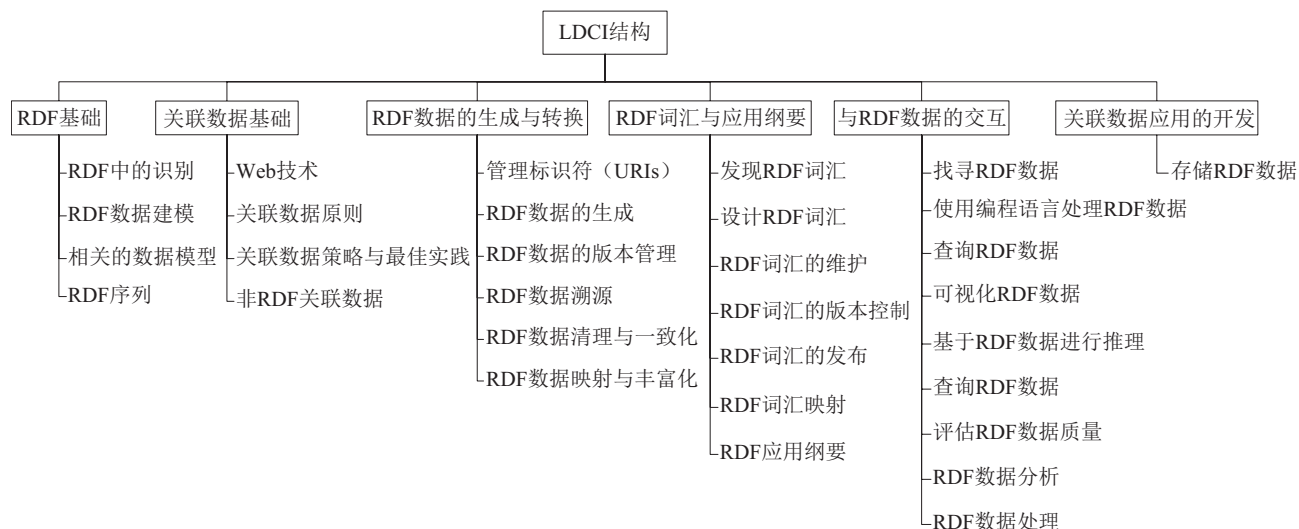


图2 LDCI两级结构展开图

待进一步扩充。

LD4PE后续运行维护主要由DCMI负责，LDCI的动态发展采用如前所述的两级合作的社区驱动模式。LDCI的编辑组制定统一的LDCI开发语言风格，作为开发准则，详细规定如何描述能力与基准的7条要求^[22]，并且明确规定能力在理解程度上分为4层，对应基准要求的技术实践操作有4种，如表2所示。

表2 LDCI语言风格对理解与实践程度等级的规定

程 度	能力（理论理解）	基准（技术实践）
低	理解（Understands）	使用（Uses）
↑	了解（Knows）	表达（Expresses）
↓	识别（Recognizes）	演示（Demonstrates）
高	区分（Differentiates）	转换（Converts）

LDCI作为关联数据学习能力知识框架具有以下特点。

（1）确定基本的学习主题范围并稳中求变。信息技术发展更迭快速，关联数据知识与技术手段也在不断发展。但是，关联数据学习的基本知识面是确定的，且相对稳定。确定基本的学习主题范围是根基，持续界定“基础”也是维系可持续发展的重要保障。

（2）采用去序列化与去核心化的方式组织不同学习主题。LD4PE学习者多样，学科背景、能力基础和认知水平复杂而不一致，这也就意味所谓的学习核心与高级是相对的。通过不同学习者实际的学习路线对比来理解，选取平台上地图页面中展示的图书馆馆员与数据科学家的学习路线进行比较^[23]，如图3所示。图中使用

序号标识能力的层级并标注于节点之中，节点对应的能力标注在节点一侧，两条路线中相同学习能力标识为同心圆。

在关联数据应用领域，图书馆馆员与数据科学家是密切关注关联数据的两类典型学习群体。图书馆馆员需将传统书目录转化为关联数据发布方式，而数据科学家侧重处理数据集格式并挖掘数据资源中的价值。在图3中，路线上的节点是有序的，但从节点能力的序号来看，LDCI中主题与主题、能力与能力之间并无学习先后的顺序逻辑。从同心圆分布来看，相同能力在图书馆馆员学习路线上相对分散，在数据科学家学习路线上集中于前半段。由此可见，不同学习群体对同一能力的掌握诉求也不同，能力之间不存在绝对的高级与低级之分。图书馆馆员关注RDF数据生成，而数据科学家侧重查询与格式转换的操作，两者的学习核心内容与侧重各不相同。以上举例可知，LDCI设计者解除了对知识框架的核心与学习逻辑的顺序设定，让关联数据学习者能够根据自身需求决定学习内容与学习路线。

2.3 关联数据学习资源的描述与组织

关联数据学习资源以LDCI为基础进行组织，学习资源与能力陈述之间的对应关系是多对多。学习资源由两部分构成：一部分是经项目组专家讨论，与资源合作伙伴共同提供，这部分资源对标能力指标学习所需的重要基础性资源；另一部分来自学习者使用学习资源元数据倡导（Learning Resource Metadata Initiative, LRMI）编辑工具上传。

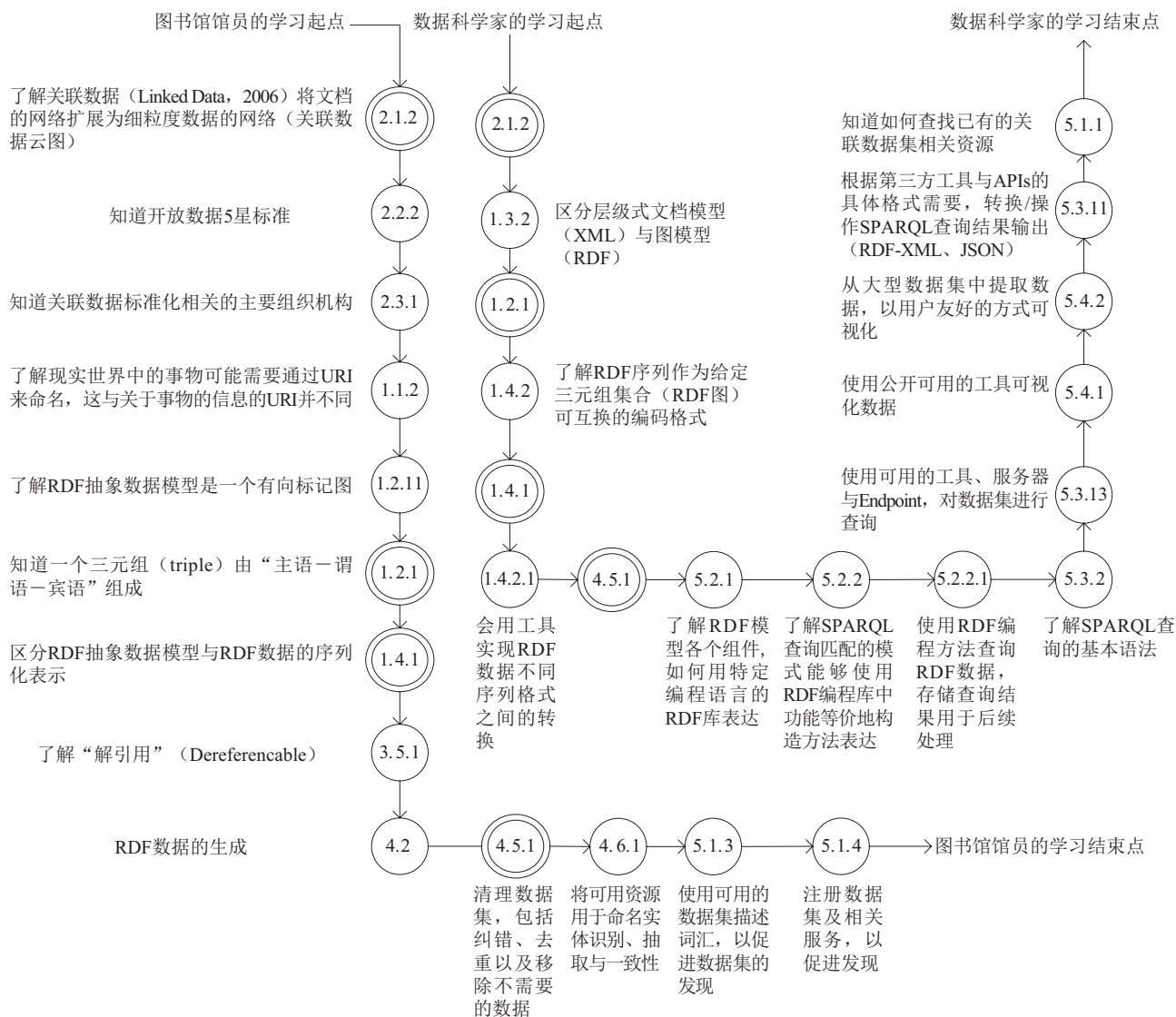


图3 图书馆员与数据科学家的学习路线对比图

表3整理总结了学习资源的通用描述、版权信息与教育相关的元数据描述。由于LD4PE复用了LRMI的元数据描述方案, 学习资源的描述包含一些通用的教育相关元数据。通过教育相关元数据 (教学用途、资源类型、资源受众、掌握程度及交互类型等) 实现多元化分类导航, 帮助学习者根据自身情况聚集相关的学习资源。基于LDCI的资源组织方式与多维度资源分类逻辑, 使得不同主题、不同难度、不同粒度的学习资源最大限度地贴合能力学习要求, 形成一种表面上无序, 实则多向关联的学习资源网络。用户在上传学习资源时, 通过LRMI编辑工具的表单对学习资源进行描述, 提交后存储在LD4PE后台的RDF存储引擎 (Triplestore)。

学习平台只给出学习资源的URL, 并不提供学习

资源本身。这样做的理由是, 不受学习资源版权问题的困扰。但是, 一个不利因素是, 由于没有稳定URL的策略, 学习资源面临URL失效的风险。目前, 平台上学习资源总计有675条, 其中项目组贡献621条, 涵盖之前整理的4类学习资源, 其中来自YouTube的视频类资源150项, Slideshare的演示文稿类PPT资源42项。LD4PE学习资源的访问权限分为免费开放和商业授权两类, 供关联数据学习者随时取用。

2.4 探究式学习功能

LD4PE平台支持探究式学习的功能主要体现在创建学习路线与保存学习资源、实践操作、学习分享与互

表3 LD4PE学习资源元数据描述

类 型	元数据元素	元数据含义	取值类型	说 明
唯一识别	资源URL*	用于学习资源的定位	URL	必备、唯一取值
通用	资源名称*	学习资源的名称	自由文本	允许多个取值
	资源描述*	学习资源的内容简介	自由文本	允许多个取值
	资源关键词*	学习资源的关键词	自由文本	允许多个取值
	资源主题	学习资源对应的能力主题	LDCI主题	例如, RDF中的识别
	资源语种*	学习资源的语种	语种代码	例如, “zh”表示中文
版权	作者*	学习资源的创建者	自由文本	允许多个取值
	出版机构*	发布该学习资源的机构名称	自由文本	允许多个取值
	创建时间*	学习资源的创建日期	日期	可选项, 格式为“年/月/日”
教育	教学用途*	特定教学场合的学习资源用途	考核、指导、专业发展	允许多个取值
	资源类型	学习资源的各种存在形式	课程、教材、大纲、数据集、学习活动等	允许多个取值
	资源受众*	能够掌握一定水平内容的资源学习者类型	学生、教师、导师等	允许多个取值
	学习时长*	典型资源受众学习该资源时所花费的平均时长	时长	例如, “P2H30M”表示平均学习2小时30分钟
	掌握程度	学习者对学习资源的掌握程度	启蒙、入门、中级、高级、专家等	允许多个取值
	交互类型*	学习资源与学习者的交互类型	互动型、说明型、混合型	允许多个取值
能力分类	资源相关的能力*	资源对应的能力名称	LDCI能力名称	允许多个取值; 资源与能力之间是多对多

注: *代表学习资源发布页面显示的主要内容

动3个方面。

(1) 创建学习路线与保存学习资源。创建个性化专属学习路线与自主选择能力对应的学习资源是LD4PE的独特之处。在LDCI基础之上, 用户通过挑选、排序与保存操作, 创建属于自己的学习路线, 赋予针对性的能力权重与逻辑顺序。上述两种活动借助Learning Map Creator与Saved Set Creator两个工具, 为学习路线或所存资源集命名, 填写内容简介并勾选是否公开, 即可完成创建学习路线或资源集。目前经典的学习路线面向图书馆馆员、数据科学家、美术馆员、网络应用开发者等学习群体。

个人所创建的学习路线, 可以跨ASN的能力指标体系进行个性化能力组合, 以连接个人原有的知识网络。例如, 图书馆馆员可以组合ASN中的图书馆领域的能力指标体系^[24]与LDCI两个指标体系中的能力, 在原有的知识网络上拓展关联数据知识框架。当大量用户组建个人学习路径, 不同背景的学习者路线将构成交叉的“交通路线”, 重要的核心能力会成为访问“客流量”巨大的站点, 客流量巨大的学习路线成为适合特定背景的学习者路线。LD4PE的探究性和普适性得以充分

体现。

在资源保存方面, 学习者依据自身的接受能力、知识侧重点与学习偏好对同一能力的学习资源进行选择。如初学者更愿意选择内容浅显的简易教程, 领域专家更偏向于艰深的技术细节的文章, 系统化学习者偏好厚重书籍与在线学习课程, 而碎片化的学习者更喜欢在线博文与PPT等轻量学习资源。这样设计不仅满足不同领域学习者的个性化需求, 还支持不同能力层次学习者进行开拓式学习。

(2) 关联数据实践操作。能力的学习是探究的开始。学习者受需求驱动进行关联数据的一系列实践, 才是关联数据能力提高的主要活动。LD4PE将学习者对能力掌握程度细分为5个等级, 即启蒙、入门、中级、高级、专家^[25]。启蒙阶段的学习者刚有概念意识, 适合于输入式学习; 入门阶段在习得理论的基础上有少许实践经验, 适合实验性的学习; 中级水平的人已能半独立地实施关联数据应用, 注重技术成功的实施与获取应用指导; 高级水平的人已能独立实施关联数据应用, 注重实际项目的开发; 专家则是被认证的权威, 注重战略研究, 指导关联数据的发展与创新。水平的区分在于实践

能力的高低。在深入学习与实践探究的过程中,同一能力会在不同难度的学习资源中得到反复学习。只有当学习者对某项能力的知识和实践经验足够丰富,才能实现能力中所要求的实践。

LD4PE还提供示例数据集作为平台的探究实验室工具,并附以操作指南文档。示例数据集由OCLC提供,来自WorldCat的书目数据。该数据集包含来自世界各地图书馆、博物馆、档案馆的书目资源,数据文件为N-Triples格式,文件总存储大小为3.6GB,三元组共计27 294 518条^[26]。示例数据集具有良好的实时性与丰富性,达到一定程度的大数据量级,为关联数据查询操作提供数据支撑。值得思考的是,LD4PE对实践能力的要求有多个层次,但现有的实践工具仅限于提供基础数据资源,对于初级学习者而言,实际操作仍存在门槛,需要搭建运行环境,对于没有编程经验的学习者而言会造成障碍。

(3) 关联数据学习分享与互动。学习平台没有为用户提供类似论坛的交流版块。学习者之间的互动主要体现为4种行为。①公开学习路线,为同行提供指导。图书馆馆员会倾向于浏览学习路径名称里包含图书馆字眼的学习路线。②评价资源。用户可通过简易的评星级给资源打分,或者进行文字评价给他人提供使用建议,但目前此功能被关闭。③通过第三方应用分享资源。学习者可通过社交软件、邮件等第三方应用分享选中的资源,将学习社区范围扩大到自身的社交网络。④在GitHub上LD4PE的仓库空间进行交流。通过拉取与请求两种操作分享个人对LDCI的发展建议。当前的学习互动效果并不理想,已公开的学习路线数目有限,参与资源评价的人数较少,未能起到学习过程中互动交流的真实效果。

从LD4PE的详细分析可以看出,LD4PE很好地回应了探究式关联数据学习场景五要素,体现了自主化、个性化与实践型等探究式学习特征,值得推广与进一步开发利用。

3 关联数据学习平台建设的本地化发展展望

关联数据探究式学习场景需要与之匹配的关联数据学习平台。关联数据学习平台的建设有助于普及关联数据教育,提升关联数据学习效能,促进关联数据资源的开发与利用。

如何在各类优质的关联数据学习资源基础上,充分利用LD4PE项目成果,探索开展关联数据学习平台的中文本地化建设是当前国内关联数据学习与培训的一个重要突破口。关联数据学习平台可辅助信息管理类专业信息组织、元数据、数字图书馆、语义网、数字人文、知识图谱等相关课程教学。

在国内信息资源管理领域,关联数据理念目前已被普遍接受,关联数据方法与技术成为信息组织的必备知识板块。但国内关联数据的系统化教学与培训尚未形成,一方面,由于关联数据的理论掌握与应用实践相对分离,学习过程中产生的技术畏难情绪、持续性问题、趣味性缺乏、收获感低等都是影响因素;另一方面,中文环境中开放的关联数据集数量有限,导致可实际接触与操作的真实数据有限,这有待国内信息资源机构加强开放资源观念,推进自有资源关联数据发布的常态化与持续化。

国内关联数据应用实践开展还不充分,目前只有上海图书馆等少数信息资源机构在积极并持续推进馆藏资源的关联数据发布。未来更进一步参与国际LOD-LAM和OpenGLAM社区交流,推动文化遗产资源的关联数据化示范应用,同时也为关联数据学习实践创造良好的资源环境。

支持探究式学习场景的关联数据学习平台本地化建设应当注重中文关联数据学习资源的建设,在保留LD4PE探究式学习设计精髓的基础上,提供更加友好的实践实验环境与技术社区氛围。目前,LDCI的中文版已经发布^[27],以LD4PE开源技术系统为基础的关联数据学习平台建设也在推进中,与之配套的关联数据实验环境通过开源的RDF存储引擎和SPARQL服务器联合打造。这将是一个复杂的系统工程,需要关联数据相关组织机构的产学研通力配合,从而实现持续性推进。

参考文献

- [1] Linked Open Data Cloud [EB/OL]. [2019-11-20]. <https://lod-cloud.net>.
- [2] International Linked Data Survey [EB/OL]. [2019-11-20]. <https://www.oclc.org/research/themes/data-science/linkedata/linked-data-survey.html>.
- [3] SMITH-YOSHIMURA K. Analysis of 2018 International Linked Data Survey for Implementers [EB/OL]. [2019-11-11]. <https://journal.code4lib.org/articles/13867>.

- [4] 王志军, 陈丽. 联通主义学习理论及其最新进展 [J]. 开放教育研究, 2014, 20 (5): 11-28.
- [5] 王妍莉, 杨改学, 王娟, 等. 基于内容分析法的非正式学习国内研究综述 [J]. 远程教育杂志, 2011, 29 (4): 71-76.
- [6] 探究性学习的理论与实验研究课题组. 探究式学习: 含义、特征及核心要素 [J]. 教育研究, 2001 (12): 52-56.
- [7] 查建中. 论“做中学”战略下的CDIO模式 [J]. 高等工程教育研究, 2008 (3): 1-6, 9.
- [8] 沈志宏, 张晓林. 关联数据及其应用现状综述 [J]. 现代图书情报技术, 2010 (11): 1-9.
- [9] 武法提. 论目标导向的网络学习环境设计 [J]. 电化教育研究, 2013, 30 (7): 40-46.
- [10] Semantic University [EB/OL]. [2019-11-10]. <https://www.cambridgesemantics.com/blog/semantic-university>.
- [11] TALIS T H. Linked Data: Evolving the Web into a Global Data Space [EB/OL]. [2019-11-10]. <http://linkeddatatoolkit.com/editions/1.0/>.
- [12] DUCARME B. Learning SPARQL: 2nd Edition [M]. Boston: O'Reilly Media, 2013.
- [13] ZAIMAN M, WOOD D, RUTH L, et al. Linked Data: Structured data on the Web [M]. Shelter Island: Manning Publications, 2013.
- [14] Linked Data School LinDa [EB/OL]. [2019-11-20]. <http://linda.seco.cs.aalto.fi>.
- [15] Library Juice Academy [EB/OL]. [2019-11-20]. <https://libraryjuiceacademy.com/054-ontologies-linked.php>.
- [16] Library Linked Data Incubator Group: Use Cases [EB/OL]. [2019-11-20]. <https://www.w3.org/2005/Incubator/ld/XGR-ld-usecase-20111025/>.
- [17] OCLC Linked Data Research [EB/OL]. [2019-11-20]. <https://www.oclc.org/research/themes/data-science/linkeddata.html>.
- [18] EUCLID-Educational Curriculum for the usage of Linked Data [EB/OL]. [2019-11-20]. <http://euclid-project.eu>.
- [19] Linked Data for Professional Education-Learning resources tagged by competency [EB/OL]. [2019-11-20]. <http://explore.dublincore.net/>.
- [20] LD4PE Project History [EB/OL]. [2019-11-20]. <http://explore.dublincore.net/about/ld4pe-project-history>.
- [21] Achievement Standards Network [EB/OL]. [2019-11-20]. <http://www.achievementstandards.org/>.
- [22] Writing effective competencies and benchmarks [EB/OL]. [2019-11-20]. <https://dcmi.github.io/ldci/style>.
- [23] LD4PE Learning Maps [EB/OL]. [2019-11-10]. <http://explore.dublincore.net/explore-learning-resources-by-competency/learning-maps/learning-map-builder>.
- [24] Competency Index for the Library Field [EB/OL]. [2019-11-20]. <http://asn.jesandco.org/resources/D2527589>.
- [25] Proficiency-levels [EB/OL]. [2019-11-20]. <http://explore.dublincore.net/theory/proficiency-levels>.
- [26] OCLC Dataset [EB/OL]. [2019-11-20]. <http://explore.dublincore.net/related/oclc-dataset>.
- [27] 范炜译. 关联数据能力指标 [EB/OL]. [2019-11-20]. <https://dcmi.github.io/ldci-zh>.

作者简介

范炜, 男, 1981年生, 博士, 副教授, 通信作者, 研究方向: 信息组织与检索, E-mail: fanw@scu.edu.cn。
张丽, 女, 1999年生, 信息管理与信息系统专业2017级本科生。

Developing a Linked Data Learning Platform in Exploratory Context

FAN Wei ZHANG Li

(School of Public Administration, Sichuan University, Chengdu 610064, China)

Abstract: The development of linked data learning platform will help popularize linked data education, improve the effectiveness of linked data learning, and promote the use of linked data resources. Guided by development and utilization, this paper explores the breakthrough of single linear learning ideas for learning problems such as complex knowledge system, relatively scattered learning resources, relatively high technical threshold, and gap between theoretical learning and technical practice. It proposes a linked data learning scenario, analyses of the scene components and the learning resource base, and in-depth analyses of the LD4PE project which is the current best practice with competency index, description of learning resource and exploratory study function. The development of linked data learning platform is a complex and continuous system engineering. Focusing on self-built learning routes, self-selected learning resource combinations, assisting practical operations, and encouraging learning to share interactions, it's necessary to develop localization of linked data learning platform based on LD4PE in the Chinese environment.

Keywords: Linked Data; Exploratory Learning Context; Learning Platform; LD4PE

(收稿日期: 2019-11-16)