

语义网教育应用研究新进展:关联数据视角

吴鹏飞, 余胜泉

(北京师范大学 教育学部 教育技术学院, 北京 100875)

[摘要] 近年来随着以关联数据为代表的语义网技术及其在政府、图书馆、媒体、企业等社会各领域中的广泛应用, 教育领域中的关联数据应用研究也引起了众多 e-Learning 研究团队的关注。文章从学习内容知识组织、学习环境创建、语义社会网络构建、学习资源整合共享、学习知识库构建等五个方面梳理国外关联数据教育应用研究进展, 综述主要研究趋势和应用实践, 最后对目前研究中存在的问题进行总结, 以期能为国内语义网教育应用研究带来一些参考和借鉴。

[关键词] 关联数据; 语义网; e-Learning; 教育; 学习资源整合

[中图分类号] G434 [文献标志码] A

[作者简介] 吴鹏飞(1982—), 男, 山东兖州人。馆员, 博士研究生。主要从事语义技术、知识本体和网络学习资源设计与开发等研究。E-mail:wupengfei_2000@163.com。

网络环境下的学习资源的共享、重用和互操作问题一直以来都是 e-Learning 领域研究的关注焦点。特别是近年来, 随着 OER(开放教育资源)和 MOOCs(大规模开放在线课程)不断发展, 学习资源日益丰富, 但是网络学习资源的混乱无序、独占隔离、简单重复、缺乏共享、检索低效等问题^[1], 使其不能满足泛在学习、移动学习、非正式学习等新的学习需求。在解决跨学习环境资源的共享与互操作问题上, 虽然有许多复杂的学习资源元数据标准, 例如 ADL SCORM、IEEE LOM、IMS LD 等, 但是在实际学习环境下还是片段化和小范围的共享和利用, 主要关注点还是学习资源的元数据创建和环境信息交换^[2]。对于 Web 范围的海量学习资源的数据互操作可以采用各种 APIs、Web Service 例如 OAI-PMH、SQI 等, 但是在个性化、情境性和智能适应性的学习资源和服务提供方面仍面临许多难以解决的挑战。

语义网强调对网络资源附加可供机器理解和处理的语义信息, 使资源的大规模重用和自动处理成为可能, 是实现未来智慧网络的数据基础, 由于语义网技术在知识组织、重用、共享和智能推理上的优势, 已引起了 e-Learning 领域研究者的广泛关注^[3]。其中语义网中

的本体和推理两大核心技术在 e-Learning 领域的应用最为广泛, 主要集中在基于本体设计和开发新型的自适应 e-Learning 系统^[4]、基于本体的学习资源组织管理与共享^{[5][6]}、泛在学习环境学习资源组织模型^{[7][8]}、基于本体和推理的个性化适应性学习内容推荐^{[9][10]}、学习资源与学习服务的语义检索^[11]、智能答疑^[12]等方面。上述研究主要是利用本体这种知识组织和表征技术对学习资源、学习者、学习服务、学习环境等进行语义化描述, 在此基础上利用各种推理技术实现 e-Learning 资源的智能适应性推荐和用户的个性化情境性学习, 研究的视角主要还是集中在面向特定教育应用的重量级本体构建和推理应用层面。另外, 重量级的本体构建是一项费时费力的工作, 需要领域专家和计算机专家协同构建, 这是制约语义网大规模应用的瓶颈所在, 加之教育领域中的概念和关系是经常变化的, 因此在教育领域中构建重量级本体也是非常困难的。

关联数据^[13]是语义网的一种轻量级实现形式, 教育领域中的关联数据应用也已经引起了众多研究团队的关注和实践, 例如英国开放大学、加拿大阿萨巴斯卡大学、德国 L3S 研究中心等, 并且出现了多个关联数据教育应用项目^[14], 如英国教育领域的 SemTech

基金项目: 教育部博士点基金项目“泛在学习环境下的学习资源进化研究”(编号: 20110003110029)

性的工作场所学习环境,该学习环境通过整合不同的 Web2.0 工具如 MediaWiki、Elgg、Twitter、Tagging Tool 等来支持职业工作场所环境下的非正式学习,利用关联数据模型表征和关联学习资源,并开发了 Learn-B 平台^[21],如图 1 所示。

三、基于关联数据的语义社会网络构建

语义社会网络是语义知识网络和社会人际网络的结合,如何在学习领域构建语义社会网络是一个前沿性的研究。对于学习领域的语义社会网络构建问题,国内北京师范大学余胜泉教授 2009 年就提出了泛在学习环境下的学习资源组织模型——学习元^[22],学习元是物化资源和人际资源相融合的资源,是知识网络与社会网络的无缝叠加,学习元体现了学习资源的社会性,是对学习领域中语义社会网络的探索与实践。关联数据对于语义社会网络构建提供了知识网络和人际网络关联机制支持,国外一些学者也开始探索基于关联数据的语义社会网络构建问题。

Kaldoudi 从语义社会网络角度探讨了教育资源的社会性问题^[23]。开发实现了医学教育资源共享平台 MetaMorphosis+, 该平台利用语义技术提升以教育资源为中心的社会性和数据整合,通过利用关联数据云建立学习资源的语义链接,并扩展和完善学习资源的社会性纬度,关联数据云的信息可以自动地丰富学习资源、扩展词汇,并创建领域特定描述和扩展医学学习资源的社会性关系。

Robinson 等提出了基于语义标注分析的解决方案,即利用关联数据的丰富描述和语义扩展为学习提供社会性纬度^[24],能够帮助学习者在相同的学习活动中通过学习其他学习者标注信息或者与其他学习者交互,为其提供兴趣最相关的学习者。基于语义标注知识库的电子书阅读能给学生提供无缝的方式去关联学习伙伴,并形成关于某个特定学习主题的同步的学习社区,大大提升了社会化学习效果。

Dimitrova 等针对在社会学习空间中不仅需要不同层次贡献者中聚合数字学习与培训经验知识,还需要提供与个性化学习与培训需求相关的学习内容问题,提出了一个利用语义丰富的关联数据支持的内容整合框架,利用了开放关联数据 DBPedia 和 SentiWordNet 本体,开发了一个多层行为本体模型 AMOn 来描述社会行为中的典型概念和关系,同时基于 AMOn 开发实现了内容智能聚合平台 I-CAW^[25]。通过利用 I-CAW 平台支持跨文化辅导和医疗面试教育应用表明,将博客、案例视频、YouTube 培训视频资

源和这些视频的相关评论等社会性学习资源以语义丰富的关联数据形式进行整合可以很好地支持非正式学习。

关联数据对于语义社会网络构建提供了知识网络和人际网络关联机制支持,解决了社会网络与知识网络相对分离的问题,基于关联数据的语义社会网络构建研究探索了 Web2.0 社会人际网络和 Web3.0 语义知识网络相结合的优势,能够实现人际关联和资源关联,改善了非正式学习环境下的学习体验,同时也提升了社会化学习效果。

四、异构学习资源整合共享

传统的学习资源组织一般采用分类体系和元数据,这种资源组织方式静态、固化,不利于扩展,同时建立在传统学习资源组织方式上的学习资源整合的粒度、层次也是不理想的,这给通过 Web 发布的学习资源的检索与利用带来了巨大挑战。关联数据具有轻量级、动态性、细粒度知识组织和跨平台、开放性、易扩展语义关联特点,在异构学习资源整合方面具有明显优势,许多研究者开始研究基于关联数据的学习资源整合与共享,出现了许多具体的实践案例。由于研究学习资源整合的侧重点不同,可将其分为三个维度:方法维度、应用维度、项目维度。

(一)学习资源整合的方法维度

Giordano 等分析了传统分类(Taxonomies)、大众分类(Folksonomies)各自的优缺点和二者之间的整合优势,针对学习资源类型和媒体类型的分类问题,提出了一种新的基于关联数据对学习资源类型和媒体类型进行分类的受控词汇开发方法^[26],用于实现医学资源的整合与共享。

Sicilia 等研究了关联数据在 Organic.Edunet 学习资源库中的应用^[27],具体包括利用 RDF 和多本体来组织学习资源,将关系数据库映射转换为 RDF 数据,利用 DCMI/IEEE draft 发布关联数据,对词条关系可视化展示,目的是将学习资源库中的学习资源发布为关联数据,从而实现跨库语义检索与浏览。

Hover 等针对在学习过程中的讨论和社会性交互活动对多媒体学习资源的细粒度标注需求,构建了一个轻量级的基于 OWL 关联数据选择器 LDS 本体^[28],用于标注和参引多媒体学习资源,为学习资源赋予语义信息,便于多媒体学习资源的整合与利用。

Siehndel 等针对微课资源中关键概念存在参考标注缺失问题,将关联数据应用到微课的语义标注中^[29],利用标注工具 Wikipedia-Miner 将可汗学院微课中的

关键概念和 Wikipedia 关联数据集中的文档进行关联,通过概念丰富性关联能够使得学习者提升对一个学习主题的相关概念的深入理解,帮助教师发现课程中被忽略的关键概念。

(二)学习资源整合的应用维度

学习资源整合应用是最终目标,来自英国开放大学、意大利巴勒莫大学等组织、机构的研究者们开始探索基于关联数据的学习资源整合应用问题。

Heath 等针对高校机构中的学习资源整合问题,设计了一个基于关联数据的教育图^[30],用于表征处于学习过程中心位置的资源的交互关系,利用基于关联数据原则和技术开发的 Talis Aspire 平台在英国开放大学中开展了基于学习资源教育图的应用,实现了学习资源实体抽取与跨数据源一致性检验,并在教育图中实现课程学习资源推荐。

Sarker 等提出了一种新的基于从机构内部和开放关联数据 HEFCE3、ONS、Unistates 中获得的可信的数据生成学生发展预测模型的方法^[31],验证了关联数据在学习数据整合服务中的优势。Bratsas 等针对传统的学习管理系统绝大部分利用学习对象技术和标准难以应用到 Web 级学习资源整合问题,在塞萨洛尼亚里士多德大学里利用维基平台 MediaWiki 和内容管理系统 Drupal 进行个性化定制与开发,将大学里的学习资源发布成关联数据为 Web 级学习资源整合、共享和利用提供便利和途径^[32]。

(三)学习资源整合的项目维度

研究领域也出现了多个面向学习资源整合应用的关联数据研究项目。如英国开放大学开展的 LUCERO 项目^[33],意大利巴勒莫大学的 LOD.CS.UNIPA 关联数据研究项目^[34]和法国高等教育机构开展的 SemUnit 项目^[35]等。

学习资源的整合共享是关联数据教育应用的重点研究内容。关联数据技术为实现 Web 范围学习资源整合与利用提供了整合方法和技术支持,关联数据在异构学习资源的整合共享方面发挥了关键性的作用。从欧洲各国高校和教育研究机构中基于关联数据的学习资源整合应用实践看,关联数据可以为教育者和学习者提供学习资源的浏览、发现与推荐等语义服务支持,为开放学习提供了便利条件。

五、开放关联学习知识库构建

学习知识库可以为学习者的个性化、情境性自适应推荐提供支持。传统的学习知识库一般采用知识工程的方法构建本地知识库,利用推理规则实现推理,

为学习提供智能服务。这种知识库建设模式存在知识库封闭、知识源单一等问题,所建的知识库一般都是限定领域的、面向教育特定应用,很难实现大规模的共享和利用。

Celino 等介绍了一个支持模拟学习系统的关联知识库构建经验与方法^[36],基于 GeoNames、Freebase、DBpedia 关联数据集整合构建了知识库,并以联邦集成 RDF 存储,知识库采用 RESTful 服务和本体进行推理,并设计了一个仿真教学系统框架,将其应用到了危机管理培训场景中。Taibi 等提出了一种利用关联数据、相似度聚类、关联和消歧技术生成大规模教育数据集的方法^[37]。Foulonneau 等研究了关联数据云作为学习资源,其可以为教育领域提供可重用的知识库,为预测学习评价难度提供了支持^[38]。

Foulonneau 等研究利用开放关联数据集生成教育评价项目,设计了一个评价自动生成平台^[39],具体生成过程包括:首先根据 IMS-QTI (IMS Question & Test Interoperability Specification, 问题与测验互操作说明规范)标准设计问题模版;其次,将问题模版中的问题与答案变量转换为标准的 SPARQL 语句,利用 DBpedia SPARQL 语义检索接口,从 DBpedia 中获取结构化语义数据;根据问题模板和语义检索返回的结果创建并生成简单的选择题模型;利用开源语义平台 TAO 发布。

Bratsas 等设计了一个利用 Greek DBpedia 作为知识库来创建网络游戏学习的案例,开发了一个游戏学习平台^[40]。该平台包括三大引擎:事实引擎是利用关联数据集 DBpedia SPARQL 端点来生成事实,共生成了 8 个类别的事实,其中包括地理、历史、政治、化学、经济等类别;查询引擎随机从生成的 8 个类别的事实库中选择一个事实并抽取出事项目传递到测验引擎;测验引擎提供用户游戏问题界面并收集用户反馈。该游戏应用到了希腊北部基础教育领域中进行了试验,结果显示该学习游戏能够充分激发学生的学习积极性和沉浸感。

Damljanovic 等设计了一个基于关联数据的智能个性化学习助手^[41],该学习助手利用知识库中的预先定义模式,经过基于 WordNet 和相关领域本体的概念选择、问题生成、答案生成、练习顺序计划生成、奖励计划生成过程,智能化生成练习,并能够根据具体情境提示、回答、激发和引导学习者理解练习题中的概念。

基于关联数据构建的开放关联学习知识库能够利用互联网环境下各个领域关联数据资源作为知识数据来源,并提供基于 SPARQL 标准的语义检索和

知识数据获取方式,解决了传统知识库建设中存在知识库封闭、知识源单一、限定领域、面向教育特定应用等问题,实现了互联网环境下知识共享和重用,为基于开放关联学习知识库的学习应用等提供知识数据基础和环境支撑。

六、总结与讨论

从上述研究和实践看,关联数据在 e-Learning 中应用具有三大优势:一是免费开放 Web 教育资源的关联数据发布可以为创新教育应用和实践提供资源、数据和知识基础;二是基于本体驱动结构化、语义化的 RDF 关联数据格式为大规模机器自动处理提供技术保障;三是跨机构的教育学习资源的语义关联和知识聚合能为学习者智能化、情境性、自适应和个性化地获取、检索和利用教育学习资源提供新的知识环境支撑。

另外,我们从教育语义维基应用、学习环境创建、语义社会网络构建、学习资源整合共享、学习知识库构建、学习评价、教育游戏设计等这些新的教育应用形式中可以看出,关联数据教育应用已经深入到了 e-Learning 中的各个方面,为 e-Learning 学习环境中的各种要素的设计、组织、管理和评价带来了新的变化。

尽管关联数据教育应用研究在近几年来取得了大量的成果,在实际的教育教学过程中也开展了很多新的实践探索,但是关联数据教育应用的发展仍然存在一些问题,值得我们进一步思考。

念及其关系描述,同时也是关联数据的核心组成要素和语义共享模型。目前全球范围已经出现了许多成熟本体如 Courseware(课件本体)、FOAF(朋友本体)、LOCO(学习对象情境本体)等。但是,目前在教育领域开发的本体是比较零散的,主要是针对特定教育应用,还没有用来描述学习者、学习活动、学习工具、学习社区、学习交互、学科知识等内容的比较成熟的和较为完善的本体。教育领域的本体构建是教育语义关联数据的语义检索、知识共享和智能化的教学应用的关键,教育领域的本体构建与利用问题是需要重点关注的问题。对于教育领域的本体构建和利用问题,一是尽可能重用共享全球范围内各领域已建立的成熟的本体词汇,二是要开发教育领域的相关本体。

(三)关联数据教育应用效果评价问题

关联数据在教育语义维基应用、学习环境创建、语义社会网络构建、学习资源整合共享、学习知识库构建、学习评价、教育游戏设计等方面都有不同的应用,不同的教育应用会带来不同的学习效果。从目前的研究现状看,对于关联数据教育应用效果评价研究还不深入和全面,对关联数据产生的学习效果的定性、定量评价还不完善。

(四)关联数据本身的动态更新机制还需完善

关联数据是一个巨大的知识网络,学习过程是一个动态的、不断变化的过程,在学习过程中,人与知识及人与人是不断交互的,因此人与知识、知识与知识以

- [2] Dietze, S., Alonso, S.S., Ebner, H., et al. Interlinking Educational Resources and the Web of Data: A Survey of Challenges and Approaches[J]. Program: Electronic Library and Information Systems, 2013,47(1):60~91.
- [3] 杨现民,余胜泉.学习元平台的语义技术架构及其应用研究[J].现代远程教育研究,2014,(1):89~99.
- [4] 曹乐静,刘晓强.基于本体和 Web 服务的适应性 e-Learning 系统[J].计算机系统应用,2005,(4):16~23.
- [5] 李艳燕.基于语义的学习资源管理及利用[D].北京:中国科学院计算技术研究所,2005.
- [6] 刘革平,赵嫦花.基于形式化本体的数字化学习资源共享技术研究[J].西南师范大学学报(自然科学版),2009,34(6):204~207.
- [7] [22] 余胜泉,杨现民,程罡.泛在学习环境中的学习资源设计与共享——‘学习元’的理念与结构[J].开放教育研究,2009,15(1):47~53.
- [8] 程罡.泛在学习环境下的学习资源共享模型——学习元的体系结构和运行环境研究[D].北京:北京师范大学,2009.
- [9] Bouzeghoub, A., Defude, B., Duitama, JF., et al. A Knowledge-Based Approach to Describe An Adapt Learning Objects[J]. International Journal on e-Learning, 2006,5(1):95~102.
- [10] 陈敏,余胜泉,杨现民.泛在学习的内容个性化推荐模型设计——以“学习元”平台为例[J].现代教育技术,2011,21(6):13~18.
- [11] 姜少峰,朱群雄. Bayesian 推理在远程答疑专家系统中的应用[J].北京化工大学学报,2003,30(6):95~98.
- [12] 郑耿忠.基于范例推理的智能答疑的研究与实现[J].微计算机信息,2008,24(12):273~275.
- [13] Berners-Lee T. Linked Data[EB/OL]. [2012-07-01]. <http://www.w3.org/DesignIssues/LinkedData.html>.
- [14] 吴鹏飞,马凤娟.国外关联数据教育应用项目研究与启示[J].电化教育研究,2013,(4):114~120.
- [15] Linked Learning 2011 e-Learning Approaches for the Linked Data Age[EB/OL]. [2013-12-01]. <http://ceur-ws.org/Vol-717/>.
- [16] LILE2013 [EB/OL]. [2013-12-01]. <http://www2013.org/papers/companion.htm#5>.
- [17] Bratsas, C., Dimou, A., Alexiadis, G., et al. Educational Semantic Wikis in the Linked Data Age: the Case of MSc Web Science Program at Aristotle University of Thessaloniki [EB/OL]. [2014-01-06]. <http://stefandietze.files.wordpress.com/2011/05/bratsas-web-science-semantic-wiki.pdf>.
- [18] Rutledge, L., Oostenrijk, R.. Applying and Extending Semantic Wikis for Semantic Web Courses [EB/OL]. [2014-01-06]. <http://stefandietze.files.wordpress.com/2011/05/semanticwikisll11.pdf>.
- [19] Vidal, J.I., Fernández, A., Díaz, A.. Thinking Semantic Wikis as Learning Object Repositories[EB/OL]. [2014-01-06]. <http://ceur-ws.org/Vol-840/10-paper-27.pdf>.
- [20] Deursen, D.V., Jacques, I., Wannemacker, S.D., et al. A Mobile and Adaptive Language Learning Environment Based on Linked Data[EB/OL]. [2014-01-06]. <http://stefandietze.files.wordpress.com/2011/05/presentatieigorjacq-ues.pdf>.
- [21] Siadaty, M., Jovanovic, J., Gasevic, D., et al. Semantic Web and Linked Learning to Support Workplace Learning[EB/OL]. [2014-01-06]. <http://ceur-ws.org/Vol-840/06-paper-29.pdf>.
- [22] Kaldoudi, E., Dovrolis, N., Giordano D., et al. Educational Resources as Social Objects in Semantic Social Networks[EB/OL]. [2014-01-06]. http://stefandietze.files.wordpress.com/2011/05/kaldoudi_dovrolis_ll2011.pdf.
- [24] Robinson, J., Stan, J., Ribiere, M.. Using Linked Data to Reduce Learning Latency for e-Book Readers [EB/OL]. [2014-01-06]. http://stefandietze.files.wordpress.com/2011/05/2011-05-29_linkedlearning_sbook.pdf.
- [25] Dimitrova, V., Thakker, D., Lau, L.. Aggregating Digital Traces into A Semantic-Enriched Data Cloud for Informal Learning[EB/OL]. [2014-01-06]. <http://ceur-ws.org/Vol-840/07-paper-30.pdf>.
- [26] Giordano, D., Kavasidis, I., Spampinato, C., et al. Developing Controlled Vocabularies for Educational Resources Sharing: A Case Study[EB/OL]. [2014-01-06]. <http://stefandietze.files.wordpress.com/2011/05/giordano-vocabulari-es.pdf>.
- [27] Sicilia, M.A., Ebner, H., Alonso, S.S., et al. Navigating Learning Resources through Linked Data: A Preliminary Report on the Re-Design of Organic.Edunet [EB/OL]. [2014-01-06]. <http://stefandietze.files.wordpress.com/2011/05/-sanchezalonso-presentation-linkedlearning2011.pdf>.
- [28] Hover, K.M., Mühlhauser, M.. Linked Data Selectors[EB/OL]. [2014-01-06]. <http://www2013.org/companion/-p439.pdf>.
- [29] Siehndel, P., Kawase, R., Hadgu, A.T., et al. Finding Relevant Missing References in Learning Courses [EB/OL]. [2014-01-06]. <http://www2013.org/companion/p425.pdf>.
- [30] Heath, T., Singer, R., Shabir, N., et al. Assembling and Applying An Education Graph Based on Learning Resources in

