

语义慕课： 语义网环境下MOOC的发展愿景*

吴文涛^{1,2}, 张舒予¹

(1.南京师范大学 教育科学学院, 江苏 南京 210097; 2.安徽师范大学 教育科学学院, 安徽 芜湖 241000)

摘要: 语义慕课又称作LOOC(Linked Open Online Courses), 即“关联开放在线课程”。近来, 国外推崇语义慕课的声音越来越大, 相关的试验应用也已生根发芽, 语义慕课已成为国际在线课程研究的一个重要方向。因此, 对语义慕课相关研究进行系统考察能为国内研究者及时跟踪国外最新动态与开展语义慕课相关研究提供借鉴。语义慕课是语义网技术与MOOC的结合体, 是一种基于MOOC平台的语义应用, 具有关联性、开放性、个性化以及智能化等特征。该文在分析德国达姆施塔特工业大学研究成果的基础上, 详细介绍了语义慕课涉及的关键技术, 并尝试构建了语义慕课的结构体系。语义慕课的关键技术主要有: 用于描述数据语义的RDF、用于构造和推理的OWL以及用于查询关联数据的SPARQL。该文构建的语义慕课完整体系主要包括四大模块: 用于搜索“货源”的数据获取模块, 用于打造“地基”的数据转换模块, 用于构建“仓库”的数据关联模块以及用于设计“窗口”的数据应用模块。未来, 我国语义慕课的研究与实践需从资源建设、局部试验、技术迁移等路径协同推进。在推进的过程中, 我们应积极面对知识产权、技术开发以及队伍建设等方面的诸多挑战。

关键词: 语义慕课; 语义网技术; MOOC; LOOC; 在线课程

中图分类号: G434 **文献标识码:** A

2008年以来, 国内外MOOC的发展大体上历经兴起、繁荣、反思以及转型四个阶段, 逐步走向“后MOOC”时代, 一批旨在优化当前MOOC的在线课程范式(如SPOC、MOOL、MOOR、PMOOC等)纷纷涌现^[1], 这些范式不仅试图从理论上为在线教育重构合用范式, 也期待在实践中给网络学习用户提供优质服务, 为在线课程的可持续发展注入新鲜血液与创新动力。转型发展过程中, 源自计算机科学领域的语义网技术受到国际在线课程领域学者的关注, 语义慕课(Semantic MOOC)应运而生。近年来, 部分国际在线课程学者就语义慕课展开研究, 取得了一系列颇具开创性的研究成果, 在国际上引起较大的关注。初识语义慕课, 我们不禁产生以下疑问: 语义慕课是什么, 它表现出了什么样的特征; 它涉及哪些技术, 又呈现怎样的结构体系。本研究拟就上述问题进行探析, 希冀为我国MOOC研究者及时跟踪国外MOOC研究领域最新动态、探索语义网环境下在线课程的设计思路提供参考。

一、语义慕课的内涵与特征

语义慕课萌生于现实困境, 肇始于技术迁移。2010年以来, MOOC的飞速发展使得主流平台中的学习资源“井喷式”涌现, 一定程度上导致学习者陷入面对海量学习资源却难以有效利用的尴尬境地。因此, 如何使得MOOC平台中的无数学习资源发挥最具意义的教育效用, 成为如今在线教育界关注的重要问题之一。与此同时, 计算机领域的语义网技术日趋完善并在部分领域初显实效。由于语义网技术能够有效解决当前网络资源存在的孤立隔离、检索困难以及共享与交互程度低等问题, 因而它被国外学者引入在线教育领域, 成为优化MOOC平台的重要技术之一。2013年, Sergio Miranda等人率先尝试运用语义网技术准确发现MOOC平台中用户的学习数据, 以优化其教学评价设计^[2]。2014年3月, Nelson Piedra等人指出, 基于语义网技术的MOOC能够提高学习者发现学习资源的可能性, 促进开放教育资源的有效利用^[3]。同年7月, 在

* 本文系2016年江苏省普通高校学术学位研究生创新计划项目“语义网环境下MOOC学习资源库的设计研究”(项目编号: KYLX16_0707)研究成果。

IEEE(Institute of Electrical and Electronics Engineers, 电气和电子工程师协会)第14届高级学习技术国际会议上,来自德国著名理工高校达姆施塔特工业大学计算机系的Kai Michael Höver与Max Mühlhäuser正式提出语义慕课这一概念,视其为语义网环境下MOOC的发展愿景^[4]。简而言之,语义慕课=语义网技术+MOOC,二者的结合赋予了语义慕课全新的内涵与特征。

(一)语义慕课的内涵解析:一种关联MOOC平台的语义应用

语义慕课又称作LOOC(Linked Open Online Courses),英文直译为“关联开放在线课程”。与一般的MOOC平台不同,语义慕课既不制作资源也不发布资源,它实际上是一种以提升现有MOOC平台中学习资源的应用效率为目的的语义应用。具体而言,语义慕课是在现有MOOC平台间建立起链接,使得他们之间能够更好地共享与交互,从而为学习者有效应用其中的学习资源提供智能化的服务。其中,语义网技术的关键作用在于为MOOC平台中的学习资源贴上能被计算机识别的标签,进而使资源自动关联在一起,并通过计算机的自动推理实现资源的智能化处理,如学习资源的精准检索、智能推荐等。不难发现,国际在线课程领域推出语义慕课的初衷便是借助语义网技术促进学习资源从孤立隔离走向互联共享,从而更大程度上发挥当前MOOC平台的教育效用。

事实上,语义慕课更像是针对MOOC学习资源的智能“淘宝网”。一方面,语义慕课的资源检索与淘宝网的搜索引擎类似,搜索引擎从无数的商品网页中抓取信息呈现给用户,而语义慕课则从大量的MOOC平台中找到学习资源提供给学习者。不同的是,搜索引擎面对的是数以亿计的松散且无序的数据,因而它如今会因检索效率底下而陷入瓶颈;而语义慕课搜索的是经过改造的具有语义关联性的数据库,它能凭借计算机的智能推理实现精准检索。另一方面,语义慕课的资源推荐类似于淘宝网的货品推荐,区别在于,货品推荐是通过大量的用户浏览记录进行数据挖掘,根据数据分析结果进行预测,而在语义慕课中,主要依靠计算机对关联数据的智能推理实现学习资源的智能推荐。

显然,语义网相关技术是语义慕课的基石。语义网技术是当前人工智能领域研究的热点,指的是一系列有关语义网的技术标准,大体上可分为三大类:语义理解技术、语义推理技术以及以前两者为基础的语义搜索技术。语义表示计算机对某一个概念、术语或者符号可以理解的“含义”^[5],语义

网技术的关键便在于以统一的规范为数据赋予这种“含义”,进而实现一系列智能化的自动操作,为用户提供个性化与智能化的服务。在语义慕课中,实现这种智能化的功能一般需要完成三个步骤:一是以语义理解技术为规范为学习资源背后的数据赋予语义;二是利用语义推理技术在有含义的数据间建立关联;三是借助语义搜索技术使这些具有语义关联的数据发挥效用。譬如,在语义慕课中,当学习者检索“爱课程平台中有哪些语义技术课程”,计算机便可理解学习者的意图是搜索课程,且范围是爱课程网,要点是语义技术,从而执行的操作是从特定数据源中准确查询到爱课程网上的语义技术课程资源并呈现给学习者。在这里,特定数据源正是通过语义理解技术与语义推理技术构建的关联数据库,计算机的智能检索则是依赖语义搜索技术来完成。

综上所述,语义网技术的诸多功能恰能解决当前在线课程领域存在的现实问题,于是,它便自然地部分学者引入在线课程的设计中,语义慕课由此而生。当前,由于语义网技术整个的技术体系仍有待完善,部分技术难题也有待攻克,因此,国外语义慕课研究也还处在理论构想与局部试验阶段。不过,这并不妨碍我们对这种智能化的课程范式进行解读与构思,为设计语义网环境下的在线课程模式进行开拓性的探索。

(二)语义慕课的特征分析

语义慕课是语义网相关技术支持下的智能教育应用,语义网技术与MOOC的深度融合使得语义慕课的特征得以凸显。虽然语义慕课仍停留于理论探索阶段,但从目前国外相关研究进展与试验应用的情况来看,语义慕课至少已展现出关联性、开放性、个性化和智能化四大特征。

1. 关联性

关联性不仅是语义网技术的核心,也是语义慕课的关键。应该说,语义慕课的关联性特征与生俱来。在语义慕课的另一名称LOOC(Linked Open Online Courses)中,“Linked”本意为“链接的”,在计算机领域就多译为“关联的”。具体而言,语义慕课的关联性有三个层面的涵义:一是数据间的连接,语义理解技术的重要任务便是在MOOC网站中的数据间建立起语义关联;二是资源间的共享,关联数据的重要作用便在于促进学习资源从孤立隔离走向互联共享;三是平台间的交互,语义慕课的最终目标是在关联数据的基础上打通现有MOOC平台间的应用壁垒,使得它们为学习者提供更加优质的服务。

2.开放性

语义慕课能够有效提升开放教育资源的开放性。2012年,联合国教科文组织将开放教育资源定义为“以数字化或其他格式存在于任何公共领域或者是根据开放许可协议发布,允许他人不受限制或者受到有限限制的免费获取、使用、改编和再分配的教学、学习和研究资源”^[6]。依此考察,现阶段的部分MOOC平台显然难以称得上“开放”,因为学习用户仅能在线学习而无法通过下载对学习资源进行其他方式的利用。相反,语义慕课不但能够促进对开放教育资源的管理、交换以及整合,而且其技术架构为开放教育资源的交换、分配和重用提供了极大的便利^[7]。正因如此,语义慕课被国际在线课程领域视作推动开放教育运动持续发展的重要课程样式之一。

3.个性化

所谓个性化,是指语义慕课能为每个学习者提供针对性的学习资源和学习服务,使得真正的个性化学习成为可能。首先,语义慕课能够为学习者提供个性化的学习资源。借助语义技术对学习资源进行多维度聚合,能够使得海量、分散、无序的学习资源聚合成能满足个性化学习的资源包^[8]。其次,语义慕课能够实现个性化的学习分析。在当前MOOC平台中,学习者的学习数据会因为多种原因而处于分散无序的状态,而语义技术能够使得计算机自动整理这些分散无序的数据,进而发现数据的所有者,从而完成个性化的学习分析。总之,语义慕课能提供个性化的学习资源,因而,个性化的学习方式才有可能发生;语义慕课能进行个性化的学习分析,由此,个性化的学习评价便有望实现。

4.智能化

语义慕课的智能化特征与个性化特征相伴而行。换言之,语义慕课为学习者提供的一切个性化服务都是以计算机智能处理为基础。离开语义慕课的智能化特征,个性化服务便无从谈起。首先,了解学习者的个性化需求需要计算机的智能语义识别;其次,为学习者提供个性化服务依赖计算机的智能语义推理。在智能化的语义慕课中,学习者要做的仅是通过一句话或者多回合的对话方式将需求告知计算机,接下来便是等待计算机提供智能化的服

务,如个性化学习资源的智能推荐、学习数据的智能分析以及新增数据的智能关联等。

二、语义慕课的技术与体系

目前,国际语义慕课研究已取得一定进展,相关经验值得国内借鉴参考。2014年,达姆施塔特工业大学计算机系的Michael Höver与Max Mühlhäuse在提出语义慕课概念的同时,也构建出“一种基于网络教育数据的语义慕课结构”(如图1所示)。他们指出,语义慕课的设计需借助三大语义网技术,即用于描述数据语义的RDF(Resource Description Framework)、用于构造和推理的OWL(Web Ontology Language,简称OWL)以及用于查询关联数据的SPARQL(Simple Protocol and RDF Query Language,简称SPARQL)。如图1所示,Michael Höver与Max Mühlhäuse搭建的语义慕课的基本结构分为三层,由下至上依次为数据网络层、数据层以及应用层。数据网络层展示了语义慕课的数据来源,数据层呈现了数据的存储方式,应用层则设计了具体的应用服务^[9]。

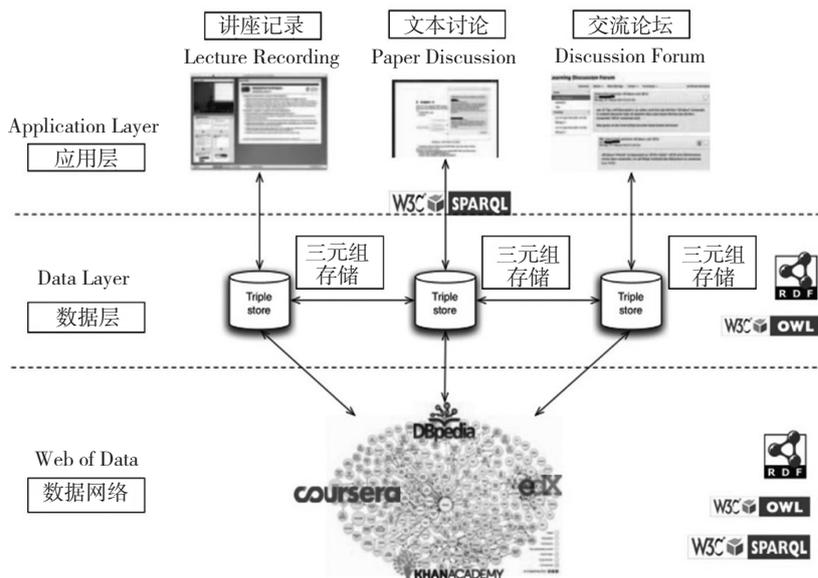


图1 一种基于网络教育数据的语义慕课结构

显然,达姆施塔特工业大学的最新研究成果兼具开创性与指导性。一方面,该研究创新的将语义网技术迁移至MOOC平台的优化设计中;另一方面,其搭建的语义慕课结构符合当前主流语义应用的基本架构,不仅对今后的语义慕课研究具有极大的指导意义,也为未来语义慕课的发展勾画出一幅令人憧憬的理想蓝图。

不过,在笔者看来,对于初识语义慕课的国

内MOOC研究学者来说,达姆施塔特工业大学的研究仍有两点稍显不足:其一,该研究主要侧重于描述应用层中具体应用的设计,未就语义慕课的基石——语义网技术的具体功能作详细阐述;其二,笔者考察诸多语义应用项目后发现,该研究搭建的语义慕课结构仅是一种简化模式,并未展示出语义慕课体系的全貌。有鉴于此,为使国内学者更为清晰地认知语义慕课的结构与功能,本研究便在以上成果的基础上做了进一步探索,主要完成了两方面的工作:一是深入解析语义慕课中三大语义网技术的具体功能;二是在对现阶段已有语义应用项目的考察的基础上,重新构建出语义慕课的完整体系。

(一)语义慕课中三大语义网技术的具体功能

如前所述,语义网技术主要分为语义理解技术、语义推理技术和语义搜索技术三类,它们具体的技术标准分别对应图1中三大关键技术,即RDF、OWL与SPARQL。目前,三大技术均由万维网联盟(又称W3C理事会)发布并形成了统一规范,是语义慕课体系中不可或缺的技术基石。

1.RDF是语义慕课中数据的描述标准

RDF是由万维网联盟发布的一种标准化的数据描述语言,它以三元组(Triple,即资源、属性和取值)作为基本数据单元来描述数据^[10]。由于现阶段MOOC平台中的数据分散且无序,要想在这些数据之间建立关联,首要任务便是以统一的标准语言描述这些数据,RDF便是当前公认的语言标准。语义慕课的设计是将当前MOOC平台中的数据均描述为RDF形式并进行存储,为后续的数据应用打下基础。

2.OWL是语义慕课中数据的关联标准。

OWL是一种被万维网联盟发布并认可的构建本体的技术标准^[11],其作用是在有语义的数据间建立起关联。在语义慕课中,只有利用OWL构建领域知识本体库,才能实现课程知识之间逻辑关系的自动或者半自动化关联标注^[12]。换言之,RDF形式的数据需要一个明确的形式规约,这样计算机才能理解数据的含义,进而准确地找到数据,发现相关的学习资源。而OWL就是这样一种通用规约,学习资源的数据依据OWL构建的本体上升为知识,可以更好地被共享和复用,从而发挥万维网的智能化潜力。

3.SPARQL是语义慕课中数据的查询标准

正如关系型数据库需要标准化的结构化查询语言来规范数据查询的表达,RDF数据也需要相应的数据存取规范。SPARQL就是万维网联盟专门为查询RDF数据开发的一种查询语言和数据获取协议,

可以用于查询任何用RDF来表示的信息资源^[13]。具体到语义慕课中,基于HTTP和简单对象访问协议的SPARQL查询能够识别被查询数据的语义及其关联,继而准确返回学习资源的数据,并以可视化网页的形式呈现给学习者。

简而言之,在语义慕课中,利用RDF可为学习资源背后的数据赋予语义;利用OWL构建本体可在语义数据间建立起关联;利用SPARQL便可查询具有语义关联的数据。三者是语义慕课实现学习资源精准检索与智能推荐的前提。

(二)语义慕课的完整体系

如图2所示,本研究所构建的语义慕课的完整体系包含四个模块,即数据获取、数据转换、数据关联和数据应用,四者各司其职,形成一种线性的逻辑关系,为语义慕课最终发挥效用提供支撑。

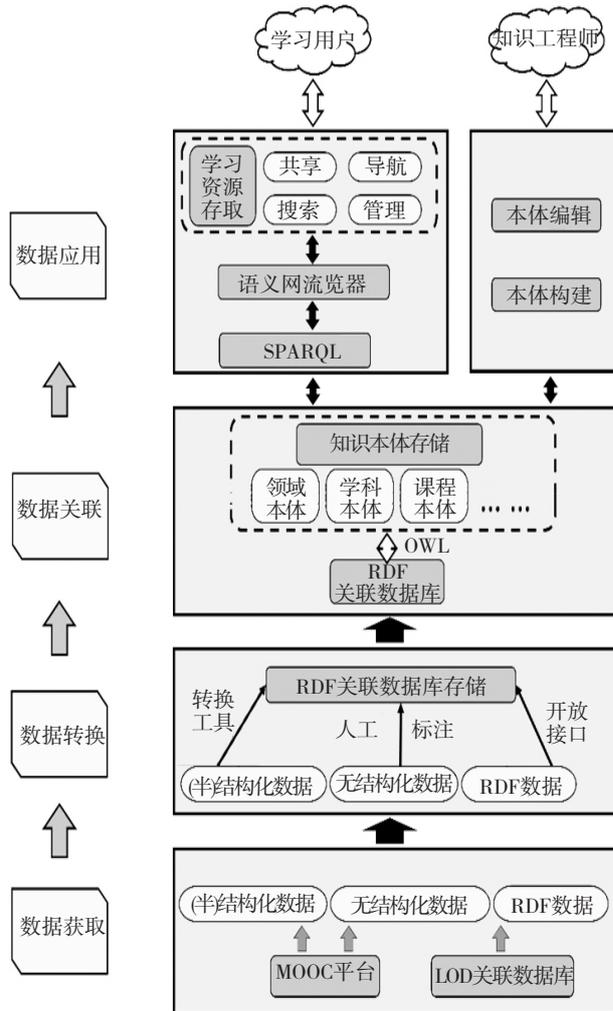


图2 语义慕课的完整体系

1.数据获取:搜索语义慕课的“货源”
数据获取模块的主要作用在于搜集学习资源

的原始数据。原始数据的来源有两大渠道：一类是主流MOOC平台，如Coursera、edX、可汗学院等，以结构化或半结构化数据为主，也会涉及少量无结构化的数据；另一类是关联开放数据集，当前，国际语义网领域最大的关联开放数据集源自著名的关联开放数据项目(Linked Open Data Project, 简称LOD)。LOD项目是由Chris Bize与Richard Cyganiak于2008年发起的倡议项目，它号召更多的个人与组织将现有网络数据发布成关联数据(Linked Data)，为语义网环境下智能应用的构建打下基础。截止2014年8月30日，已经有来自全世界的570个数据集被发布成链接数据且相互间建立起关联，DBpedia是其中最大也最为著名的数据集^[14]。近年来，这些关联数据集已经为搜索、金融、语音领域的语义应用提供了有力的数据支撑。

需要补充说明的是，从以上两大渠道获取数据各有利弊。仅从技术角度看，获取LOD关联数据库的RDF数据显然更加便捷，数据所涉及的网络资源也更加丰富，然而，由于LOD关联数据库中的数据量过于巨大且涵盖太多领域，构建知识本体将是一项浩大的工程，难度也无法估计^[15]。相比之下，主流MOOC平台的数据虽然比较有限，但是这些数据涉及的主题更为集中，从其中抽取RDF数据并构建知识本体更有望实现。正因如此，目前国外研究也多基于主流MOOC平台的数据设计各类语义应用。譬如，加拿大学者Gilbert Paquette与Alexis Miara借助语义网技术设计了一种名为COMETE智能资源管理工具，用于提升MOOC平台的智能检索能力^[16]；美国学者Sebastian Kagemann与Srividya Bansal则搭建了一个被称作MOOCLink语义网站，实现了对Coursera平台上学习课件的智能获取与对比^[17]。

2. 数据转换：打造语义慕课的“地基”

数据转换模块的主要任务是赋予数据语义。显然，有含义的数据是语义慕课的基石，缺少它，语义慕课便是无本之木、无源之水。为数据赋予语义需要将数据网络中获取的各类数据转换成RDF形式，即从原始数据中抽取三元组形式的轻量级的RDF数据并进行存储。如前所述，原始数据因来源渠道不同大体分为三类：结构化或半结构化的数据、无结构化数据以及已经关联的RDF数据。结构化或半结构化的数据多源自关系型数据库，本身已有语义内涵，仅需按照RDF形式进行统一调整；无结构化的数据由于本身缺乏语义信息，需要重新描述；已经关联的RDF数据则无需转换可直接调用。

一般认为，三类数据的转换分别有三种方式。对于结构化或半结构化的数据，可以通过运

用相关转换工具，使用浅表层分析技术抽取语义形成RDF数据。譬如，万维网联盟与麻省理工学院共同开发的转换平台SIMILE RDFizer就提供多个RDF转换器，能够将关系型数据库中的结构化或半结构化转换成RDF数据，其中和教育有关的转换器包括 MARCRDFizer、OAI-PMH RDFizer、OCWRDFizer^[18]。对于无结构化的数据一般多采用人工标注的方式抽取语义，操作相对复杂且费时费力。目前，部分转换平台(如Drupal、Fedora、Dspace、Eprints等)也能够通过统计技术和浅自然语言技术为人工标注数据提供一定的帮助。对于已经关联的RDF数据，数据库一般都会提供具有开放接口的SPARQL终端，支持其他应用直接查询与使用其中的数据。

3. 数据关联：构建语义慕课的“仓库”

数据管理模块的关键功能是依据OWL构建知识本体，在语义数据间建立起关联。足够的知识本体不仅能使得RDF数据得到有效的组织和管理；还能进一步增强RDF数据间的语义与逻辑关系，为计算机的智能推理提供支持。本体是某个领域的概念集合和概念之间的关系，通过形式化的方式表达该领域的知识及其知识结构^[19]。一个本体是一个特定领域的模型，为特定的目的而构建^[20]。在语义慕课中，由于存在多种特定的需求，往往需要构建不同类型与不同层次的知识本体，如领域本体、学科本体、课程本体等等。

语义网环境下，本体的构建方法一般有三种，即专家手动构建、计算机自动构建以及本体复用。专家手动构建本体耗时费力，已经无法满足当前的需求；计算机自动构建又存在准确率不够的弊端；而本体复用则是指充分利用当前网络中已经存在很多的知识本体以减少工作量。由于当前网络的公共领域中存在许多现成的本体，因此我们很容易找到一个可用的第三方本体，为我们构建新的知识本体提供一个有用的起点。此外，数据关联模块的另一功能是用于数据存储。前期形成的RDF数据以及后期构建的各类知识本体及其实例均需存储于其中。

4. 数据应用：设计语义慕课的“窗口”

数据应用模块是语义慕课的顶层，主要任务是查询数据、存储数据以及修改数据。在前文案例中，Michael Höver与Max Mühlhäuse也探讨了MOOCs平台中最常涉及的三种应用，即交流论坛(Discussion Form)、文本讨论(Paper Discussion)、讲座录像(Lecture Recording)。三种应用均需通过SPARQL端点实现对数据的调用。需要强调的是，面向RDF数据的查询与当前搜索引擎中的关键词查

询不同,这是一种更加智能化的语义查询。也就是说,计算机能在识别搜索关键词的语义信息的基础上自动推理,为用户呈现具有语义相关性的各类资源,使得资源精准检索与智能推荐成为可能。

考虑到一般用户的数据处理能力有限,在语义慕课中,SPARQL查询工作并不直接面向一般用户,需要经由语义网浏览器将数据连接的学习资源以可视化的方式呈现给用户。近十年来,语义网领域的多个研究机构先后推出了十几个有影响力的语义网浏览器,如麻省理工学院开发的Piggy Bank、英国开放大学开发的Magpie以及爱尔兰数字企业研究所开发的VisiNav等。以VisiNav为例,它是一个支持可视化查询构造的语义网数据访问系统,能够实现对RDF数据集的6种基本操作:关键词搜索、对象导航、分面式浏览、路径遍历、投影和排序^[21]。一般而言,基于语义网浏览器的语义慕课能够以共享、导航、搜索、管理等方式促进学习用户充分利用学习资源。此外,语义慕课也支持学习用户以某种方式(如上传、发布、记录等)为其增加新的学习资源,这些资源也将以三元组的形式为描述关联数据并存储下来,在知识本体的组织下供其他用户重复使用。

除了以上的查询与存储功能,数据应用模块还需为知识工程师提供服务。由于学习资源的动态变化以及不同用户的各种需求,知识工程师需要不断构建新的知识本体、及时修正旧的知识本体,并将它们放回到存储库中,为语义慕课的持续有效工作提供技术支持。

三、我国推进语义慕课的实施路径

理解语义慕课的内涵特征与结构体系的重要目的之一便在于促进我国语义慕课研究提上日程。语义慕课是国际在线课程领域一个全新的课题,它的推出使得当前MOOC向智能型课程平台转型成为可能。在语义慕课的支持下,学习用户能够更加开放、更加智能、更加灵活的利用学习资源,由此,更加个性化学习与智能化学习也有望实现。当前,国外推崇语义慕课的声音越来越大,相关的试验应用也已生根发芽,语义慕课已成为国际在线课程发展的一个重要方向。国际语义慕课的理论性构想与试验性研究为我国推进语义慕课研究的开展提供了宝贵的经验与启示。

(一)以优质资源为语义慕课研究提供保障

认清语义慕课的本质便能把握语义慕课的关键。套用一句知名广告语,语义慕课不生产教育资源,它只是教育资源的智能搬运工。如果说MOOC

的关键作用是将最优质的教育资源无偿提供给世界上每个角落的学习者,那么语义慕课的重大意义便是促进所有人更好地发现与使用最优质的教育资源。显然,离开优质的教育资源,语义慕课便是无本之木、无源之水。从开放课件项目到开放教育资源,再到大规模开放在线课程,优质资源建设不一直是开放教育资源运动的题中应有之义,也始终是在线课程领域致力解决的核心课题。只有优质教育资源被建设好,语义慕课才有发挥价值的可能。因此,我国在线课程要想尽快走向更加智能化的语义慕课,在突破相关的技术瓶颈之前,首要任务之一是建设优质的在线教育资源,为语义慕课发展奠定坚实的资源基础。

(二)以局部试验为语义慕课前进开拓道路

国外研究启示我们,局部试验性研究是推动语义慕课发展的重要方式之一。如前所述,原始数据是语义慕课的“货源”。因此,开展语义慕课试验研究的第一步是获取足够数量的原始数据。那么,从哪里获取合适的原始数据呢?显然,最佳的方式是从现有在线课程平台中获取。然而,由于国内很多的MOOC平台系企业开发,未能提供开放的API(Application Programming Interface,应用程序编程接口),以致后台的数据无法直接调用。因此,获取数据的可行方式之一便是依靠国家教育行政部门开发的公益性在线课程平台。2003年以来,国家教育行政部门牵头开发了一大批精品课程资源,建成了多个在线课程平台,如国家精品课程资源网等。这些平台中的数据不仅符合语义慕课研究的要求,有些还因存在共享、重用和复用率低等问题^[22]而使得开展语义慕课的试验研究具有更大的价值。因此,相关研究人员可以在征得教育主管部门同意的前提下获取相关平台中的部分数据,在此基础上,以试验性研究为国内语义慕课的前进开拓道路。

(三)以技术迁移为语义慕课发展注入动力

语义网技术源自计算机科学领域的技术革新,其应用目前也多集中于搜索引擎、金融以及语音等领域。正是一批具有超前视野的研究学者将语义网技术引入e-Learning领域,它才引起在线课程研究者的关注。应当看到,新兴技术的迁移应用不仅是在线课程领域突破的“助推器”,也是语义慕课发展的“动力源”。显然,要想使得语义慕课从愿景走向现实,相关技术瓶颈的突破是必须解决的关键问题之一。就目前研究来看,数据的语义转换技术与本体的构建技术是其中的重点之一。然而,仅凭在线教育领域的研究人员单方面的努力,突破这

些技术瓶颈难度太大。因此,必须借力于国内语义网领域的专家,以实现相关技术的顺利迁移。只有关键技术得以成功迁移,语义慕课才有可能从闭门造车走向集思广益,从单打独斗走向协同合作,从静止式孤立发展走向永动式持续前进。

四、我国语义慕课发展中需要关注的相关问题

(一)知识产权方面

“开放”一直以来都是开放教育资源运动强调的核心要义,也是语义慕课设计的关键理念。然而,随着社会对知识产权的愈发强调,落后的产权保护政策将会在一定程度上对语义慕课的发展产生阻滞。如前所述,语义慕课的关键将现有MOOC平台中的数据转换成关联数据,这是一种对第三方数据的消费和再创造,过程中必然涉及他人的利益,如何明确不同类型和归属的数据(或数据集)的所有权、发布权、使用权、收益权等问题都将是语义慕课发展过程绕不开的关口^[23]。鉴于我国知识产权保护方面仍旧依靠《中华人民共和国知识产权法》《中华人民共和国著作权法》等法律法规,尚未制定针对网络资源与开放数据的相关法律法规,部分发达国家现有的网络许可协议(如英国的开放政府许可协议)值得我国借鉴与参考。此外,由于现有大多数MOOC平台均需注册、登录方可进入,这也给语义慕课的发展带来麻烦。试想,即便是语义慕课能将各大MOOC平台学习资源的关联在一起,但学习者在利用不同平台的资源时仍需多次涉及注册、登录等环节,如此,语义慕课的效果也将大打折扣。显然,上述问题将成为语义慕课发展过程中必须积极面对与妥善解决的重要课题。

(二)技术开发方面

语义网技术的突破是语义慕课发展的基石与前提。尽管在万维网联盟的推动下,一系列技术规范得以发布与公认,语义网的理论基础已经奠定,但语义网数据集成、语义网搜索、基于语义网的智能问答和语义网软件工具等方面仍旧存在部分技术难题有待攻克。倘若以上技术瓶颈难以突破,那语义慕课便是“空中楼阁”,无法从理想走向现实。不过,当前语义网技术已在部分领域初露锋芒,语义网技术体系也正日趋完善,局部范围的有效应用已经让学界看到希望。相信伴随着时间的推移,部分技术难题也将在知识工程师们的努力下得到解决,语义慕课便有望从理论构想走向实践应用,从局部试验走向全面开展。

(三)队伍建设方面

语义慕课的发展显然离不开相关研究人员的

推动与参与。由于语义慕课不仅是一个全新的方向,也是一个涉及计算机与在线教育的交叉领域,因而目前国内从事相关研究的人员远远不够,研究队伍亟待建设。事实上,我国早就非常重视语义网技术的相关研究,国家863计划从2002年起便将其列为重点支持项目,国内清华大学、南京大学、东南大学等多所高校都设有语义网相关技术的研究机构,也产出了一批高质量的研究成果^[24]。问题的关键是,这些研究成果并没有在在线教育领域得到应用,其中的主要原因在于国内的语义网技术与在线教育研究各自为政,尚未形成有效的沟通与合作。因此,当前亟待解决问题的是如何促进语义网技术领域与在线教育领域的协同合作。显然,最好的方式之一是从以上两个领域中吸引一批研究人员,组建一个或多个研究团队,从而为推动我国语义慕课向纵深发展提供充分的人员保障。

参考文献:

- [1] 祝智庭,刘名卓.“后MOOC”时期的在线学习新样式[J].中国远程教育,2014,(10):36-43.
- [2] Miranda S,Mangione GR,Orciuli F,et al. Automatic Generation of Assessment Objects and Remedial Works for MOOCs[DB/OL].http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6671018,2016-04-25.
- [3] Piedra N,Chicaiza J,Lopez J,et al. Supporting Openness of MOOCs Contents through of an OER and OCW Framework Based on Linked Data Technologies[DB/OL].http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6826249,2016-03-23.
- [4][9] Höver KM,M ü hlhäuser M. LOOCs: Linked Open Online Courses: A Vision[DB/OL].http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6901536,2016-02-26.
- [7] Piedra N,Chicaiza J,Lopez J. An Approach for Description of Open Educational Resources Based on Semantic Technologies[DB/OL].https://www.researchgate.net/publication/224148709,2016-04-13. An_approach_for_description_of_Open_Educational_Resources_based_on_semantic_technologies. 2016-04-03.
- [5] 王继新,黄涛等.基于语义网格的教育知识服务体系研究[J].中国电化教育,2006,(5):93-96.
- [6] 保罗·川内,肖俊洪,杨伟燕.开放教育资源质量保证准则TIPS框架[J].中国远程教育,2013,(19):11-21.
- [7] Piedra N,Chicaiza J,Lopez J. An Approach for Description of Open Educational Resources Based on Semantic Technologies[DB/OL].https://www.researchgate.net/publication/224148709,2016-04-13. An_approach_for_description_of_Open_Educational_Resources_based_on_semantic_technologies. 2016-04-03.
- [8] 吴鹏飞,余胜泉.学习资源语义关联关系及其可视化研究[J].中国电化教育,2015,(12):97-104.
- [10] RDF [EB/OL].https://www.w3.org/wiki/RDF,2016-03-28.
- [11] Patel-Schneider PF, Hayes P, Horrocks I, et al. OWL Web Ontology Language Semantics and Abstract Syntax[J]. Encyclopedia of Information Science & Technology Second Edition, 2004,(1):7-15.
- [12] 余胜泉,万海鹏.支持课程大规模开放的学习技术[J].中国电化教

- 育,2014,(7):7-18.
- [13] 百度百科. SparQL[DB/OL]. <http://baike.baidu.com/view/661304.htm>,2016-04-23
- [14] The Linking Open Data cloud diagram [EB/OL]. <http://lod-cloud.net/>,2016-04-28.
- [15] 李曼,杜小勇等.语义Web环境中本体库管理系统体系结构研究[J].计算机研究与发展,2006,43(z3):39-45.
- [16] Paquette G, Miara A. Managing Open Educational Resources on the Web of Data[J]. International Journal of Advanced Computer Science & Applications,2014, (8):55-63.
- [17] Kagemann S,Bansal S.MOOLink: Building and Utilizing Linked Data from Massive Open Online Courses[DB/OL].<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7050836>,2016-04-13.
- [18] 陈大庆,丁培,叶兰,等.MOOC教育资源语义化关联研究[J].图书情报工作,2014,(14):121-126.
- [19] 丁国柱,余胜泉.基于本体学习算法的学科本体辅助构建研究——以学习元平台语文学科知识本体的构建为例[J].中国电化教育,2015,(3):81-89.
- [20] Grigoris Antoniou.语义网基础教程[M].北京:机械工业出版社,2014.
- [21] 瞿裕忠,胡伟,程龚.语义网技术体系[M].北京:科学出版社,2015.
- [22] 王娟,刘名卓,祝智庭.高校精品课程应用调查及其对精品资源共享建设的启示[J].中国电化教育,2013,(12):40-46.
- [23] 张春景,刘炜,夏翠娟等.关联数据开放应用协议[J].中国图书馆学报,2012,(1):43-48.
- [24] 胡伟,瞿裕忠,黄智生.语义Web课程建设初探[J].计算机教育,2013(12):69-72.

作者简介:

吴文涛:在读博士,讲师,研究方向为信息技术与课程整合、媒介素养教育(owen2015@qq.com)。
张舒予:教授,博士生导师,研究方向为视觉文化与媒介素养。

Semantic MOOC: The Development Prospect of MOOC under the Environment of the Semantic Web

Wu Wentao^{1,2}, Zhang Shuyu¹

(1. College of Science Education Nanjing Normal University, Nanjing Jiangsu 210097; 2. College of Science Education Anhui Normal University, Wuhu Anhui 241000)

Abstract: Semantic MOOC is also called LOOC, which is an acronym for Linked Open Online Courses. Recently, more and more attention has been drawn to Semantic MOOC, the relevant tests have been started. Semantic MOOC has become an important direction of international online course area. Therefore, in order to enable the domestic scholars to follow the international trends and to carry out related research, it is necessary to make a systematic investigation on Semantic MOOC. Semantic MOOC is the combination of semantic web technology and MOOC, and it is a kind of semantic application for linking MOOC platform. The characteristics of MOOC are related, open, personalized and intelligent. Based on the analysis of the research of Darmstadt Industrial University in German, this paper introduces the key technologies of Semantic MOOC, and tries to construct the complete structure of Semantic MOOC. The key technologies are Resource Description Framework, Web Ontology Language and Simple Protocol and RDF Query Language. The complete structure mainly includes four modules: data acquisition module(for searching data), data conversion module(for transforming data), data correlation module(for linking and storing data) and data utilization module(for using data). In the future, the research and practice of semantic MOOC should be promoted from aspects of resource construction, local experiment and technology transfer. In the process of advancing Semantic MOOC, it should positively cope with the challenges, such as problems of intellectual property rights, problems of technical development and problems of team building.

Keywords: Semantic MOOC; Semantic Web Technology; MOOC; LOOC; Online Courses

收稿日期:2016年6月14日

责任编辑:李馨 赵云建