

基于学习微元的资源共享机制研究与实践

刘涌 蒋勇

(西南科技大学 计算机科学与技术学院, 四川 绵阳 621010)

摘要: 针对目前流行的 e-learning 学习标准 SCORM 及学习管理系统体系结构在资源构建、资源共享机制方面存在的问题, 提出了一种基于学习微元的学习资源共享机制。本文首先在对学习微元的粒度和结构定义的基础上, 对学习微元的制作与发布、资源的组织与检索、课程内容聚合等过程提出了一套较为完整的解决方案, 在此基础上开发了应用原型系统。该系统能有效地降低学习资源内部耦合程度, 以主题搜索引擎模式构建学习资源, 为 e-learning 资源共享与知识融合提供开放的支撑环境, 为基于知识的学习资源集成研究提供了灵活的应用平台。

关键词: e-learning, 学习资源共享, 主题搜索引擎

随着计算机的广泛普及和网络技术的飞速发展, 远程教育技术得到了迅速推广, 基于 Web 的远程学习管理系统如雨后春笋般迅速涌现。LOM、IMS、AICC 等 e-learning 学习标准的制定为学习资源共享注入了新的活力, 由美国国防部的 ADL 计划所推出的 SCORM 2004 更是由于集成了 LOM、IMS、AICC 等成熟的研究成果而得到了广泛的支持。得益于这些标准在资源描述、资源聚合、资源检索与导航、资源运行环境等方面所提供的统一规范, 人们可以选择大量的资源制作工具制作具有丰富表现力的媒体资源(如文本、图片、视频、音频、PPT、动画)、共享和重用别人制作的教学资源、为教学资源制定灵活的课程学习计划、方便地跟踪学习者的学习过程和进度、为个性化学习提供专家系统决策指导……。纵观 e-learning 的研究进展可以发现, 可重用性、可共享性、互操作性是 e-learning 的主要研究内容。

一、前言

在深入研究这些 e-learning 标准和广泛使用过基于这些标准的学习管理系统之后, 我们发现 SCORM 标准也并不是尽如人意。在资源共享和重用性方面, 由于其体系结构方面的原因, 使得资源共享与重用形式不够灵活, 各个独立的资源制作工具以及学习管理系统之间的资源共享和重用的范围有限, 严重制约着资源提供者、计划制作者、学习管理系统和学习者之间的资源交流。具体表

现如下:

1、以资源制作者为中心的紧耦合资源聚合机制

在 SCORM 的内容聚合模型 CAM 中, 以 Asset 元数据(包括文本、图片、视频、音频、flash 等)作为资源物理文件描述的基本单位, 以 SCO 元数据(包含多个 Asset 及元数据的教学单元)作为资源调度的基本单位。在更高层次上, 使用内容聚合元数据从总体上描述内容聚合体(如内容包)以方便对内容聚合体进行检索, 使用内容组织结构元数据描述内容组织结构的详细信息, 使用活动元数据描述一个单独的学习活动以指导学习活动过程。其中, 所有元数据都是基于学习资源物理文件相对位置的资源描述信息, 必须与学习资源物理文件结合到一起才能够发挥作用。这样, 模型中的各种组成元素构成了一种紧耦合系统, 使得制作的整个学习内容聚合包缺乏任何一个方面都不能独立工作。这种模型结构完全体现了以资源制作者个人为中心的思想, 客观上将资源制作、描述、聚集的任务全部集中到资源提供者身上, 离开了资源描述元数据的资源 (Asset、SCO) 无法实现独立的共享和重用, 不利于学习计划的自由组织。

2、以 LMS 为中心的资源共享机制

综合考虑 SCORM 的三个相对独立的规范(内容聚合模型 CAM、排序与导航 SN、运行时间环境 RTE), 对学习管理系统内部的资源制作、共享与重用、互操作等方面提

供了完整可行的系统方案，但对独立支持 SCORM 规范的资源制作工具(如 PPT、Flash 等)、跨系统的多个学习管理系统之间的资源共享与重用支持甚少。按照 SCORM CAM，各系统之间的学习资源需要以 PIF 内容聚合包或者 CD 等介质进行资源交流，客观上限制了资源共享和重用的范围，这种以 LMS 为中心的资源共享机制无法提供一种有效的手段支持各类独立系统之间的资源重用与知识融合。

正是基于以上两个方面的原因，导致了参与 e-learning 的各类角色(包括资源提供者、学习计划制作者、资源制作工具、学习管理系统、学习者等)之间在共享和重用学习资源方面存在着巨大的鸿沟，不利于信息时代的资源共享与知识集成。因此，在四川省教育厅青年基金项目(知识集成的关键技术研究, RKTKI)支持下，在对 SCORM 内容聚合模型进行了深入研究的基础上，提出了构建灵活的资源基本组织单元、降低学习资源内部耦合粒度、以搜索引擎模式构建学习资源共享与重用体系结构的思想，为 e-learning 资源融合提供开放的支撑环境，为基于知识的学习资源集成研究奠定坚实的基础。

二、开放的 e-learning 资源共享模型

实现 e-learning 资源的广泛共享，需要兼顾参与 e-learning 的各类角色(包括资源提供者、学习计划制作者、资源制作工具、学习管理系统、学习者等)在资源制作、课程计划制定、课程学习进程等各个阶段的实际需求，为资源的存储、检索、集成提供一种开放、简单、高效的资源描述与聚合模型。

1、资源基本组织单元——学习微元

前已述及，SCORM 将资源基本单元(Asset 和 SCO)与内容聚合元数据、内容组织元数据、活动元数据紧耦合在一个内容聚合包中的方法不利于资源基本单元的独立共享。而在现实活动中，将资源基本单元的制作和教学活动的策划独立开来的做法更加有利于资源的共享与重用。因此，需要定义一种更加灵活的资源基本单元组织结构，该结构应该包含基本物理文件(如文本、图片、视频、音频、PPT、动画)，能实现自

我描述(如对基本物理文件的类型、所属学科、课程、模块、单元、知识点等)，并能方便地进行检索与访问。

学习资源是一种数字资源，WEB 目前已经成为数字资源的主要开放共享形式。为方便资源的检索与获取，实现灵活广泛的资源共享，可以将嵌入了基本物理文件内容的网页作为资源基本单元的组织形式。为支持基于 WEB 的资源搜索，还需要对 LA 施加一些具体的限制，定义如下：

学习微元(Learning Asset, 简称 LA)：是学习资源的基本构建块，表现为一个符合 html 标准的网页。该网页中包含了一些指向基本物理文件的链接，并能够在 WEB 中展示这些物理媒体内容；同时，网页中还必须包含对所提供内容的自描述元数据，元数据结构符合 LOM 规范，由资源提供者提供描述信息。

```
<html xmlns="http://www.w3.org/1999/xhtml">
<head profile="http://www.w3.org/2003/g/data-view" >
<link rel="schema.lom" href="http://ltsc.ieee.org/xsd/LOM" />
<link rel="transformation"
href="http://e-learningWG.swust.edu.cn/xsl/lom-extract.xsl"/>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<lom:general>
<lom:identifier>
<lom:catalog>URI</lom:catalog>
<lom:entry>http://e-learningWG.swust.edu.cn/content/co_01
</lom:entry>
</lom:identifier>
<lom:title>
<lom:string language="en">Title for the LRA</lom:string>
</lom:title>
<lom:language>en</lom:language>
.....
</lom:general>
<lom:technical>
<lom:format>
<lom:string language="en">text/html</lom:string>
</lom:format>
</lom:technical>
</head>
<body>
.....
</body>
</html>
```

图 1 学习微元 LA 结构示例

众所周知，超文本标记语言 HTML 易于理解，在目前基于 WEB 的网络应用中得到了广泛的支持，各种媒体形式的基本资源都可以嵌入 HTML 中，并通过网络浏览器展示(如文本、图片等)或调用客户端安装的播放程序播放(如 PPT、音频、视频、动画等)。同时，HTML 标准允许在其内容中插入一些资源描述信息。正是 HTML 网页具

有这两方面的特点，使得我们选择 LA 作为学习资源的基本单元成为一种可行的途径。按照 LOM 规范，学习微元可以定义为如图 1 所示的结构。

2、学习微元的制作与发布

基于以上定义，LA 的制作过程包含三个方面的内容：①使用常用的第三方工具（Power Point、Flash、Media Encoder 等）制作展示内容，并保存为基本物理文件；②将基本物理文件嵌入标准的 HTML 网页；③在网页中嵌入符合 LOM 标准的资源描述元数据。

在 LA 制作完成后，需要将 LA 上传到 WEB 服务器（在以后的描述中，我们称这类提供学习微元的 WEB 服务器为“学习微元网站”）并在 Internet 上发布出来，以支持基于 LA 的学习资源检索与访问。

就像我们经常访问一些资源网站一样，学习微元网站是一类专门支持 LA 发布的资源网站，提供对学习资源的集中存储和方便访问。

与 SCORM 的内容聚合模型相比，不难看出：由于资源的存储可以分布在网络中不同的学习微元网站，并具有了自我描述能力，因此不再需要使用内容聚合元数据和内容组织结构元数据来描述资源，从而降低了课程资源内部的耦合程度，使得学习微元可以作为一个独立的资源被灵活地存储、检索和重用。

3、基于主题网络爬虫的学习微元检索机制

在现代社会中，搜索引擎已经成为人们不可或缺的工具，搜索引擎技术为资源的共享与重用提供了一种自由开放的平台。然而，由于缺乏对资源内容的描述，使得人们往往需要花费大量的精力来筛选所想要的资源。虽然在 LA 中包含了对资源较为详细的描述信息，但通用搜索引擎的工作方式仍然不能正确提取出这些有用的信息。因此，需要针对学习微元定义一种特殊的资源搜索机制，以支持对学习微元的检索与聚集。在 PKTKI 中，我们基于主题网络爬虫提出了一种学习资源检索机制，以提供对学习资源这一特定领域的检索支持。

作为搜索引擎的基础组成部分，网络爬

虫（又称“网络蜘蛛”）起着举足轻重的作用。网络爬虫技术目前在各类搜索引擎中已经应用得非常普遍，它们随时运行在各个网站之间，为搜索引擎搜集网页资源，为资源搜索提供基础素材。在专门针对学习微元检索的开放系统中，我们采用了主题网络爬虫来实现对资源描述元数据的采集，支持学习微元的检索和聚集。

在专门针对学习微元检索的开放系统中，我们采用了一种二级服务机制——元数据采集服务和元数据聚集服务——分别实现对资源描述元数据的采集和提供学习微元的检索支持：

一方面，元数据采集服务运行在主题搜索引擎的元数据采集服务器中，通过其上运行的网络爬虫，按照链接深度搜索各学习微元网站，并提取各学习微元中包含的元数据，形成为标准的 HTML 网页，按照域名及目录结构存放到本地文件系统中，或者存储到数据库中。

另一方面，通过运行在元数据聚集服务器中的元数据聚集服务，遍历元数据采集服务器中各目录或者数据库中的记录，提取出其中的元数据，合并为 RDF 图模式，以支持基于语义的资源查询。

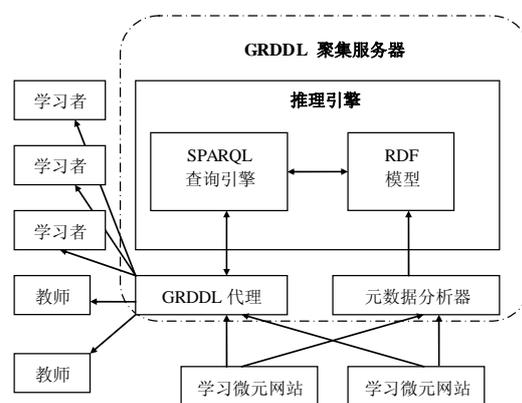


图 2 学习微元聚集与检索机制

4、基于学习微元的课程内容聚合

制定学习计划是教师的主要工作，学习资源的广泛共享其实是对教师工作的一种很好的支持。将教学活动元数据与基本组织单元元数据分开制作和存放的方法有利于提高基本组织单元元数据的利用率，支持教师针对不同教学对象聚合不同的学习课程，以引导学生的学习进程。

在构建了基于主题搜索引擎的学习微元检索机制后，课程学习计划（体现为一系列教学活动元数据）的制定过程就变得非常简单，可以以一种问答式的交互过程由学习管理系统来具体实现：

首先，学习计划制订者检索元数据聚集服务器，按照课程、模块、单元、知识点的层次结构来组织所要制定的课程计划。由于聚集服务提供了基于语义的知识查询功能，因此能够给出具体课程的知识组织体系供用户选择。

接下来，元数据聚集服务器根据用户选定的课程计划知识点层次结构检索其 RDF 推理机，逐一返回各知识点最匹配的资源元数据列表，并链接至该资源实际存放的 URL 地址。

然后，学习计划制订者逐一访问与学习计划中各知识点相对应的学习微元，以确认在学习计划中使用哪些学习微元。

最后，学习计划制订者将选定了学习微元的课程计划保存起来，或者按照 SCORM 的打包机制将课程计划和所有选定的学习微元包装成 PIF 课程包，以支持学习者的学习过程。

三、应用模式设计

按照上一节描述的开放 e-learning 学习平台结构，我们提供了一种典型的应用模式，如图 3 所示。其中包含了资源制作者、资源制作工具、学习微元网站、元数据聚集服务器、聚集服务器、计划制订者、学习者共七类参与角色。下面对各种角色在该平台中的作用及交互过程简要介绍如下：

资源制作者：资源制作者使用资源制作工具（如 Power Point、Media Encoder、Painter、Flash 等）制作符合展示要求的媒体资源，保存为对应的物理文件，上传至某个学习微元网站，并负责填写该资源的相关描述信息（如资源类型、所属课程、模块、知识点等信息）；

资源制作工具：为资源制作者提供资源素材和编辑能力。目前已有部分资源制作工具提供了对 SCORM 规范的支持，可以将制作的资源导出为 SCORM 的 Asset 或 SCO。因此，在本文所描述的平台中，我们期望资

源制作工具可以提供检索服务，在制作资源的过程中检索聚集服务器为用户提供丰富的资源素材；更深入的支持还应该包括提供对资源的描述信息的编辑，以及将制作的结果按照学习微元的标准形式进行保存。

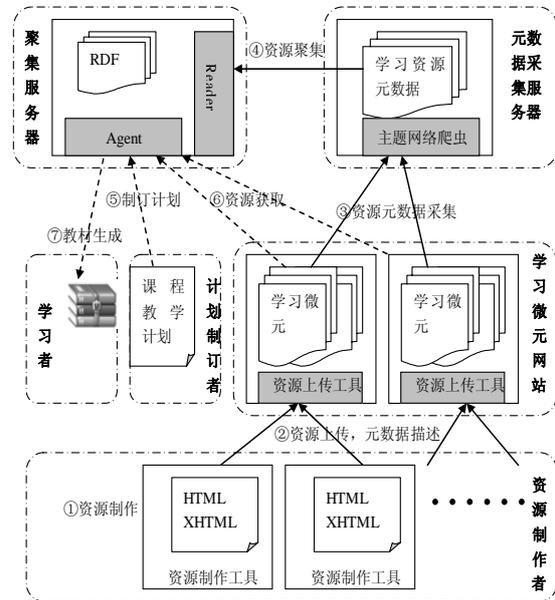


图 3 学习微元共享与聚集原理模型

学习微元网站：学习微元网站运行着 WEB 服务，为资源制作者提供资源上传功能和符合 LOM 标准的元数据填写界面，并将用户填入的内容集成到学习微元中，在其自身的 WEB 服务器上发布出来。

元数据采集服务器：元数据采集服务器与聚集服务器一起为基于学习微元的资源检索与聚集提供服务。具体说来，元数据采集服务器通过运行的主题网络爬虫，从各个学习微元网站中提取学习微元，供聚集服务器进行资源的检索和聚集。出于实际应用考虑，为了降低网络访问数据量和减少存储空间浪费，通常可以只采集学习微元中的元数据而非整个学习微元；同时，由于研究领域的特殊性，可以考虑通过在元数据采集服务器上提供注册信息，由各个学习微元网站向元数据采集服务器主动注册，为主题网络爬虫提供深度搜索起点，进一步降低网络数据访问量，节省硬件和通信方面的费用。

资源聚集服务器：资源聚集服务器从元数据采集服务器提取学习微元元数据，为计划制定者或学习者提供资源检索与聚集服务，可以集成与学习管理系统相似的功能

(如学习管理、课程管理、用户管理、测试管理、社区管理等), 以支持学习过程。为了能够向用户提供基于语义的课程计划制订, 或者具有专家系统指导的课程学习, 在聚集服务器中以支持语义表述能力的 RDF (资源描述框架) 描述学习微元元数据, 并以结合了推理机的 Agent 为资源检索、聚集和智能学习提供知识推理能力支持, 以 SPARQL (RDF 数据查询语言) 提供对 RDF 的检索服务。

计划制订者(教师): 通过聚集服务器提供的对课程知识的推理能力和课程计划制作工具, 将已有资源聚集成课程学习计划, 以类似 SCORM 活动元数据的形式保存计划, 或者按照 SCORM 标准生成 PIF 课程包, 为学习者提供学习内容和学习过程指导。

学习者(学生): 遵循课程教学计划或者教材内容开展学习。

四、结束语

在目前支持 SCORM 标准的众多学习管理系统中, 从功能上看, 无一不是围绕课程教材为教师、学生提供诸如教学活动设计、课程管理、角色管理、学习管理、作业管理、测试管理、师生交互等一体化的功能模块, 资源的共享与重用只能局限于各个学习管理系统内部, 人为地在各个学习管理系统中形成了“信息孤岛”, 不利于知识融合的研究和发展。

在按照本文所述模式构建的资源共享平台体系中, 将资源发布服务、主题搜索引擎服务引入 e-learning 学习共享平台的模式有利于打破各个独立的资源制作工具、学习管理系统之间在共享和重用资源方面的障碍, 让资源提供者、资源搜索服务提供者、课程管理与学习管理服务提供者之间都能够各司其职, 共享学习微元资源, 共同为 e-learning 学习服务。

在 PKTKI 中, 构建了第三节所描述的参考应用模式原型系统。在该原型系统中, 将开源的 weblech 改造为主题网络爬虫提供多线程、多网站学习资源搜索, 以 HP 实验室

语义 Web 项目组发布的 GRDDL Reader 和 Jena 作为基于语义的元数据聚集工具和语义推理引擎。目前已在应用原型上开展了针对计算机组成原理和计算机接口技术两门相关课程的资源服务, 基于语义匹配的资源检索和智能学习指导服务研究正在深入进行。

[参考文献]

- [1] Advanced Distributed Learning(ADL) Initiative, “SCORM 2004 3rd Edition-Impact Summary Version 1.0”. <http://www.adlnet.gov/scorm/index.aspx>, 2006.
- [2] IEEE LTSC, “The Learning Object Metadata standard”, <http://www.ieeeltsc.org>, 2002.
- [3] Broisin, Julien , Vidal, Philippe, Meire, Michael, and Duval, Erik, “Bridging the gap between learning management systems and learning object repositories: Exploiting learning context information”, Proceedings - Advanced Industrial Conference on Telecommunications/Service Assurance with Partial and Intermittent Resources Conference/E-Learning on Telecommunications Workshop AICT/SAPIR/ELETE 2005, 2005, p 478-485
- [4] AICC Management and Processes Subcommittee, “Reusability Analysis of Reusable Objects”, <http://www.aicc.org/>, 05-JUNE-2003
- [5] 刘金红, 陆余良, 主题网络爬虫研究综述, 计算机应用研究, Vol. 24, No. 10, Oct. 2007, pp.26-29.
- [6] HP Labs Semantic Web Programme, “Jena – A Semantic Web Framework for Java”, <http://jena.sourceforge.net/>
- [7] W3C, “Resource Description Framework (RDF)”, <http://www.w3.org/RDF/>, 1999.