

数字化校园设计与构建方法研究

祝伟华¹ 杨 丹^{1,2} 桑 军¹ 方蔚涛² 文俊浩¹ 徐光侠¹

(重庆大学软件学院 重庆 400044)¹ (重庆大学虎溪校区管委会 重庆 400044)²

摘 要 本文就数字化校园中软件系统建设提出了一套较为完整的、系统的具有较强可操作性的方法,应用 EAI 的各种关键技术,对该方法中各个环节要解决的问题进行了较详细的分析,并对解决这些问题的方案进行了较深入的研究和探讨。提出了以视图设计为主体的自上而下的数字化校园规划设计方法,和一种自下而上的搭建系统的方案。

关键词 数字化校园,URP,EAI

On the Design and Construction of Digital Campus Software System

ZHU Wei-Hua¹ YANG Dan^{1,2} SANG Jun¹ FANG Wei-Tao² WEN Jun-Hao¹ XU Guang-Xia¹

(College of Software,Chongqing University,Chongqing 400044)¹

(Huxi Campus Administration Boards,Chongqing University,Chongqing 400044)²

Abstract This article gives a complete, systematic and practical method for digital campus software system construction. It utilizes the key technologies in EAI area and gives detailed explanations on key problems; also it gives the solutions to solve these problems. Mainly the article proposes a top-down method to design the system and a bottom-up method to construct the system; comprehensively it gives a complete methodology and solution.

Keywords Digital campus,URP,EAI

1 引言

数字化校园是以网络为基础,利用先进的信息化手段和工具,实现从环境、资源、到活动的全部数字化,在传统校园的基础上构建一个数字空间,以拓展现实校园的时间和空间维度^[1]。

数字化校园中存在不少问题,主要表现在几个方面,一是缺乏整体规划。二是缺乏统一规划,信息不能有效共享。三是彼此之间不能实现数据访问,形成信息孤岛。四是用户登录系统需分开进行,没有形成统一的接口。五是没有统一有效的规划,整个数字化校园的扩展性较弱^[2]。

EAI 技术能较好地解决以上数字化校园中的问题。EAI 技术包括数据集成、应用集成和业务流程集成等多个方面。具体到技术层面上的划分,一套完整的 EAI 技术层次体系应该包括应用接口层、应用整合层、流程整合层和用户交互层四个大的层面。

EAI 技术可以解决数字化校园建设中的问题。但现在的数字化校园建设中往往重视技术而方法不足,注意力集中在已有的个别系统及个别软件上,普遍着力于解决已经存在的问题,头疼医头,脚疼医脚,往往只见树木不见森林,这样得到的方案和设计往往对个别子系统具有较强的依赖性,耦合度很高,生命周期短。究其原因没有一个是系统化的方法,建设中具有较大的随意性。另一方面数字化校园项目有其自身的特点,在 EAI 技术和方法的使用上,不能完全照搬,要有一定的取舍。因此,有必要形成一套行之有效可供所有数字化校园项目使用的有特色的可行方法和方案。

在重庆大学虎溪校区数字化校园建设中,采用了以两条线索为中心的基于三种视图的规划建设方法。该方法是一种系统化的较为完整的可供执行的方法。

2 数字化校园规划建设方法探讨

2.1 自上而下的规划与设计方法

数字化校园成败的一个重要因素是统一领导、统一规划、统一标准、统一管理,这是一个自上而下的过程^[3]。只有使用一个自上而下的规划与设计才能产生一个完整的、一致的具有较长生命周期的方案,才能使该方案不依赖于现有系统,从长期来看才是经济的和有效的。

2.2 自下而上的集成与搭建方法

运用 EAI 技术在校园中与企业中是有较大差别的,企业 EAI 是基于资金流及产品流的,企业的各个部门关系非常密切,可以采用整体规划下的紧密耦合的体系结构来构建。而针对校园 EAI 来讲,大学具有非集中式的校园文化,各部门有很强的独立性,是一种松耦合方式,不宜使用 ERP 建设中自上而下的建设方法。另一方面,长期以来由于学校各职能部门在信息化中投入了大量的人力、物力,各种业务严重依赖现有的系统,无法抛弃已有系统从头开始。因此在全局上,自上而下的规划与设计后,在实施上需采取自下而上的集成方法,充分利用已有资源来搭建系统。

2.3 具体的三种视图方法

自上而下的规划与设计,具体可以采取基于数据视图、服务视图、流程视图的设计方法,在实施和构建中采取相关 EAI 集成技术来实现搭建这三种视图。三种视图和已有系统间的关系如图 1 所示。

3 基于视图的设计及构建方法

3.1 数据视图方法

3.1.1 数据视图设计方法 数据视图方法主要是为数据集成和数据共享服务。数据集成不是对某个部门数据的简单

整理,而是跨越学校各个应用系统,在全校信息模型和数据视图的基础上,构建一个合理的数据存储模式。建立跨部门的数据交换平台,使学校各个应用系统具有合理的数据分布并

满足各个应用系统之间的数据交换需求,满足为全校教师员工、学生提供全方位信息服务的需求,满足学校决策的信息支持需求,并为服务视图和流程视图的构建打下基础。

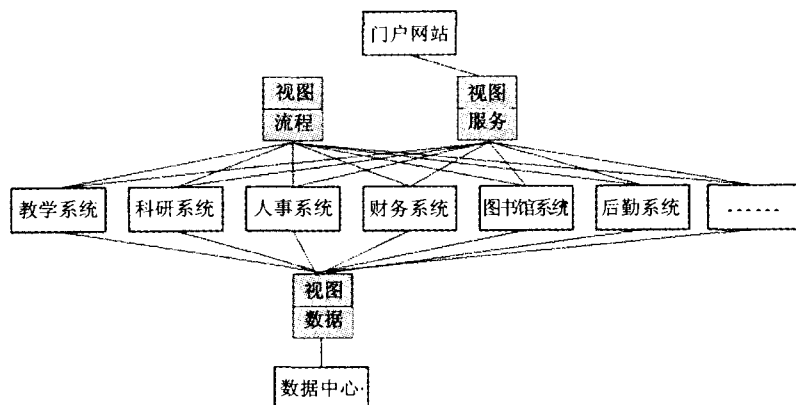


图1 数字化校园视图方法

数据视图的设计应在广泛业务调研基础上,进行实体建模,建立数据中心视图。其中一般需定义以下要素:(1)实体;(2)实体间关系;(3)实体的拥有者;(4)实体的使用对象及各种权限;(5)实体数据的来源;(6)实体数据的规范及定义等。

一方面该数据视图并不依赖已有的应用,视图和应用可以独立开发并升级。另一方面统一的数据视图又可以作为一个数据的权威,数据中心将内容推送到相应的应用系统,以达成数据共享的目的。

从应用的角度可以将结构化数据分为以下三种,即:(1)私有数据:业务部门产生的供自己系统使用的数据;(2)部门交换数据;(3)公共数据:为多数系统提供的共享数据。

数据在数据中心的交换数据库和公共数据库服务器加以存储。数据的共享与交换通过数据共享交换平台实现,数据共享交换平台由共享数据库、数据交换系统和全局数据视图组成。通过数据交换系统实现各业务子系统之间的数据交换与共享,真正实现一个业务系统数据存储对其他业务系统的透明,使业务系统之间的耦合度降到最低。各业务子系统的私有数据存放在各部门本地系统。

实施中要定义具体的数据库、帐号、访问方式、数据映射关系等,兼顾到已有的应用及数据,自下而上的将已有的数据从各个子系统中映射到统一的数据视图中达成数据集成和共享。

3.1.2 基于复制的数据视图构建方法 可以采用应用数据库系统中的异步数据复制、同步数据复制、数据联邦等手段来达成数据间的集成^[4],在建立数据视图前必须先制定数据编码标准,数据标准,数据权限规范,其中数据权限规范描述了数据的来源,数据共享的范围与方式。

(1)数据复制的方法。数据复制分为同步复制和异步复制。复制的需求分为几类:一是往其他数据库分发数据;二是整合其他数据库的数据到数据中心;三是与其他数据库进行双向数据交换。

在复制过程中,程序将数据抽取出来,通过中间的传输,最后加载到目的地。复制往往不是一次性的批量操作,需要周期性地获取数据源的增量变化信息,并将其送到目的地。复制应该不干预现在的应用,对系统的影响应该降到最低,实

现松耦合。复制过程有必要进行管理和监控。

基于学校的特点,采用异步数据复制,以数据中心为主导的数据抽取和推送的方案是可取的^[5]。

(2)数据复制中的权限方案。在异步复制环境中,对于所有应用,最关键的就是要确保数据的一致性。在同一时间对同一个表的同一行数据的同一列在两个不同的地点作更新,这种情况就会发生称之为更新冲突的错误。为保证数据的一致性,更新冲突必须被检测到并加以处理以确保在不同地点的数据元素保持同样的值。更新冲突可以通过限制“所有权”到单一节点或者将更新某个特定数据元素的权利限制到某一具体节点的方法来避免。

为确保实际应用数据的一致性,必须在异步复制应用模型中考虑冲突避免或者冲突检测和消除的方法。有两种常用的冲突避免方法:主站点所有权模型和共享所有权模型。

主站点所有权是指异步复制数据被一个单一节点“所有”,这些要复制的数据仅能被该节点更新,其他节点向拥有该数据的主节点订阅数据,这意味着它们在本地系统上只能够存取这些复制数据的只读拷贝。一旦有变更,数据中心通过中央信息的分发的方式,将这些信息复制到远程的相关业务子系统的数据库中。

共享所有权模型是通过异步的方式进行的,会遇到冲突检测和和处理问题,所以该复制方案需同时考虑冲突避免和冲突检测与消除两种情况。

简洁的主站点所有权模型是实现数字化校园中数据视图权限的一个较好选择,可以避免数据潜在冲突给业务规则和系统实施上带来的复杂性。

3.2 服务视图方法

3.2.1 服务视图设计方法 对于学校来说,各学院与行政部门可能负责很多网络设备与不同的应用系统,需要使用不同的系统、不同的界面,对不同的对象进行管理,从而导致管理人员被动地应对多套管理工具、多种形式的告警,大大降低了资源的利用效率和管理人员的工作效率,更谈不上实现学校的整体管理。

另外我们倡导一种面向服务的管理,从全校各部门提供的服务的角度来管理相关的系统。从管理系统来说,通过对业务层的抽象,可以提供给相关管理人员一个统一的服务视

图。

数字化校园的建设最终是要提供优良的服务给师生员工,是否能够提供给这些用户简捷方便的服务是成败的关键,因此提供给用户一个统一的完整的服务视图是必须的,访问该视图应该是一个唯一的入口及用户使用一个唯一的帐号。

校园的用户最主要的是教师和学生,用户数量也最大,其次是领导和管理人员用户。服务视图设计的首要任务就是明确定义每个角色所需的服务,在此基础上合成有关的服务视图。首先应解决的是主要用户学生和普通教师的服务,其次是领导和管理人员。在这个独立的完整的视图基础上,再把已有的应用系统上的服务映射过来;我们所做的服务可以是两类:基于已有应用的映射和在数据视图上新开发的应用服务。在实际的应用过程中,这两类是结合使用的,不过已有应用的映射,具有较强的耦合性,在条件允许下应考虑优先使用在数据视图上新开发的应用服务。

服务视图的设计主要有以下要素:(1)服务名称;(2)服务提供者;(3)服务的使用者;(4)服务相关的实体;(5)服务的权限;(6)服务的有关输入,输出详细定义等。

3.2.2 基于门户集成服务的构建方法 服务视图的实现体现在门户技术,单点登录上。使用一个标准的用户界面来实现应用界面的集成,满足该标准的应用界面可以聚合在该门户中;另一个方面使用单点登录和统一用户认证来达成统一的用户帐号^[6]。门户还支持通过 Web 把服务连接到各类设备上。门户网站为站点用户提供了到多种类型的信息和应用程序的单点访问和以一种适应用户的方式聚集多种类型及多种来源的信息。一个完整的门户网站解决方案使用户可以方便地访问他们完成任务所需的每一样工作。

可以利用框架的思想来构建门户,符合这些标准的应用接口都可随时地集成到该门户网站中。该框架是开放的、可扩展(称为 Portlet)的,在此基础上可以构建和部署满足多种需求的门户网站,可以支持包括机关部门对部门,学院对学院,学院对教师和学生对学院的多种校务模型。整个门户系统的开放架构,可与学院现有、以后的业务子系统进行集成,可满足学院教学工作和机构改革的动态发展需求。

门户的个性化的特点可以用它来组织属于学校自己的门户网站视图、管理学校的概要文件并发布和共享文档。门户

也可以提供附加的服务,如单点登录(Single Sign-On)、安全性、内容管理、搜索和分类、对移动设备的支持以及站点分析等,这些功能对整个学校的服务集成建设可以起到重要作用。

3.3 流程视图方法

3.3.1 流程视图设计方法 数字化校园的信息整合最高阶段就是流程整合,即利用工作流、消息、协同等技术,实现跨系统的流程整合,使得不同的管理部门能够在一个统一的网络办公环境中实现对同一事务的协同处理,为用户提供真正意义上的电子办公服务。流程整合不仅需要对应应用有较大的修改和规范,同时涉及到管理流程的再分析与重构。

通过建立一个全局的学校业务流程视图,明确定义出这些典型的流程,然后使用流程集成的方法来集成已有应用中的子流程来实现全局的流程,定义每个流程所属的使用者(角色),每类角色相应的权限。通过建立全局的流程视图,再使用流程集成的方法来将各部门的子流程进行整合,最后搭建成一个唯一的、完整的全局流程视图。通过该流程视图提供的信息,可以在网上实现以前需人工完成的相关部门的各个步骤程序。

流程视图的设计主要有以下要素:(1)流程名称;(2)流程的使用者;(3)流程的管理者;(4)流程相关的实体;(5)流程过程描述;(6)流程的权限;(7)流程的输入,输出详细定义等。

3.3.2 基于 SOA 的流程集成方法 对应流程视图方法的相关实现技术可以使用 Web Method、SOA、MQ 等。SOA (Service Oriented Architecture)即面向服务的架构是流程整合的基础架构,它是一个组件模型。它将应用程序的不同功能单元通过这些服务之间定义良好的接口和契约联系起来,通过 Web 服务的方式构建 SOA 达成流程集成。SOA 以用户为中心的原则做设计,完全按照用户的需求整合相关的组件资源,可以帮助学校面对变化的需求,使应用根据业务的需要变得更加灵活,以适应不断变化的校园环境。

结束语 综上所述,在数字化校园中可以系统化地使用自上而下的方法完成以服务视图、流程视图、数据视图为中心的设计及规划,再使用自下而上的方法基于已有的系统使用 EAI 的各种成熟技术搭建所有视图来完成系统建设。在重庆大学的虎溪校区数字化校园建设中使用了该方法,其中 URP 平台规划如图 2 所示。

重庆大学虎溪新区 URP 平台

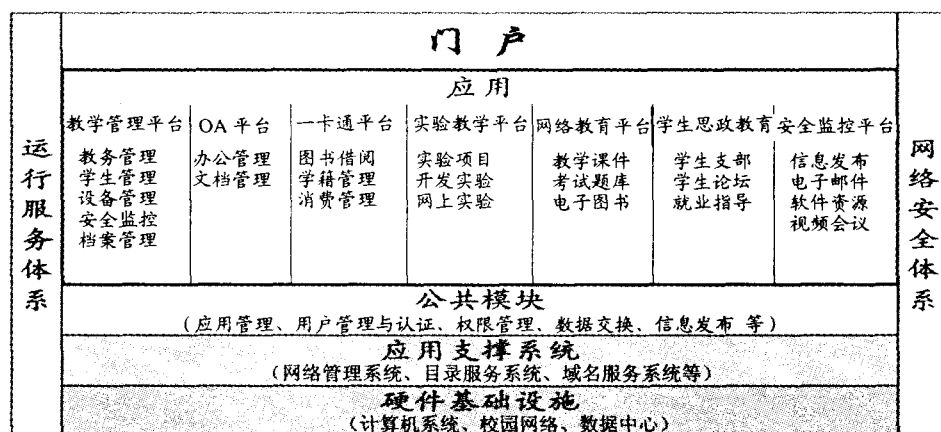


图2 重庆大学虎溪校区数字化校园模型

```

(8)  $FC_i \leftarrow FC_i \cup \{c\}$ ; //将  $c$  并入  $FC_i$  中
(9) end
(10)  $FCC_{i+1} \leftarrow \text{Gen-Generator}(FC_i)$ ; //产生  $FCC_{i+1}$  中的 generator
(11) end
(12)  $\text{Answer} \leftarrow \bigcup_{j=1}^i \{FC_j, \text{closure}, FC_j, \text{support}\}$ ;

```

算法 2 Gen-Closure 函数

```

(1) forall objects  $o \in O$  do begin
(2)  $G_0 \leftarrow \text{Subset}(FCC_i, \text{generator}, f(\{o\}))$ ; //若 generator 是  $f(\{o\})$  的子集, 将其并入  $G_0$ 
(3) forall generators  $p \in G_0$  do begin
(4) if ( $p.\text{closure} = \emptyset$ ) then
(5)  $P.\text{closure} \leftarrow \{o\}$ ;
(6) else  $p.\text{closure} \leftarrow p.\text{closure} \cup f(\{o\})$ ;
(7)  $p.\text{support}++$ 
(8) end
(9) end
(10)  $\text{Answer} \leftarrow \bigcup \{c \in FCC_i | c.\text{closure} \neq \emptyset\}$ ;

```

算法 3 Gen-Generator 函数

```

(1) insert into  $FCC_{i+1}.\text{generator}$ 
(2) select  $p.\text{item1}, p.\text{item2}, \dots, p.\text{itemi}, q.\text{itemi}$ 
(3) from  $FC_i.\text{generator } p, FC_i.\text{generator } q$ 
(4) where  $p.\text{item1} = q.\text{item1}, \dots, p.\text{itemi-1} = q.\text{itemi-1}, p.\text{itemi} < q.\text{itemi}$ ;
(5) for all generators  $p \in FCC_{i+1}.\text{generator}$  do begin
(6) for all  $i$ -subsets of  $p$  do begin
(7) If ( $s \notin FC_i.\text{generator}$ ) then
(8) delete  $p$  from  $FCC_{i+1}.\text{generator}$ ;
(9) end
(10) end
(11) for all generators  $p \in FCC_{i+1}.\text{generator}$  do begin
(12)  $S_p \leftarrow \text{Subset}(FC_i.\text{generator}, p)$ ;
(13) for all  $s \in S_p$  do begin
(14) if ( $p \subseteq s.\text{closure}$ ) then
(15) delete  $p$  from  $FCC_{i+1}.\text{generator}$ ;
(16) end
(17) end
(18)  $\text{Answer} \leftarrow \bigcup \{c \in FCC_{i+1}\}$ ;

```

4.2 生成频繁项集

下面给出由频繁闭项集导出频繁项集的算法。

算法 4

```

(1)  $k \leftarrow 0$ ;
(2) for all frequent closed itemsets  $c \in FC$  do begin
(3)  $L_{\|c\|} \leftarrow L_{\|c\|} \cup \{c\}$ ; //将频繁闭项集按照长度插入  $L_i$ 
(4) If ( $k < \|c\|$ ) then  $k \leftarrow \|c\|$ ;
(5) end
(6) for ( $i \leftarrow k; k > 1; i--$ ) do begin
(7) forall itemsets  $c \in L_i$  do begin
(8) forall ( $i-1$ )-subsets  $s$  of  $c$  do begin
(9) if ( $s \notin L_{i-1}$ ) then begin
(10)  $s.\text{support} \leftarrow c.\text{support}$ ;
(11)  $L_{i-1} \leftarrow L_{i-1} \cup \{s\}$ ; //将  $s$  插入  $L_{i-1}$  的最后
(12) end
(13) end
(14) end
(15)  $\text{Answer} \leftarrow \bigcup_{i=1}^k L_i$ ;

```

4.3 关联规则生成

关联规则生成这一部分我们仍采用 Apriori 算法中关于规则生成的算法, 这里我们不再详细论述。

4.4 实验结果

我们在某市邮区中心局物资管理信息系统中所积累的数据的基础上对上述算法进行了比较, 在已有的数据库上我们

进行了一些转换工作, 将事务数据库原始数据库的离散值进行分类和量化, 转换后的数据库中的值是布尔型的。经过转换生成的事务数据库的规模为记录数 40000 余条, 项目(属性)个数为 200 条(可以分为 6 类), 图 1 是两种算法的挖掘情况的对比图。

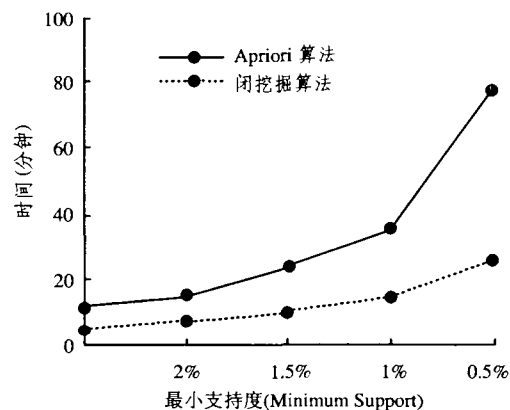


图 1 两种关联规则挖掘算法的对比图

从图中可以看到闭挖掘算法所耗用的时间比 Apriori 所耗用的时间要少, 实验结果表明闭挖掘算法是优于 Apriori 的。

结束语 本文中通过引入格论, 对关联规则的经典算法 Apriori 算法进行了改进, 此算法通过减少候选项集的方法, 既减少了寻找频繁项集的系统开销, 又减少了数据库的扫描次数。我们对这两种算法进行了对比, 实验结果表明, 算法在特定的数据库中是有效的。

参考文献

- Keim D A. Visual techniques for exploring databases. In: Tutorial Notes, 3rd Int. Conf. on Knowledge Discovery and Data Mining (KDD97), Newport Beach, CA, Aug. 1997
- 何炎祥, 等. 基于网络环境的分布式 KDD 及 Data Mining 研究. 小型微型计算机系统, 1999(10)
- 王志海, 等. 概念格上规则提取的一般算法及渐进式算法. 计算机学报, 1999(1)
- 胡长流, 宋振明. 格论基础. 河南大学出版社, 1990
- Pasquier N, Bastide Y, Taouil R, et al. Efficient Mining of Association Rules Using Closed Itemset Lattices. Information Systems, 1999, 24(1): 25~26
- 朱明, 等. 数据挖掘. 中国科学技术大学出版社, 2002
- Zaki M J, Ogihara M. Theoretical Foundations of Association Rules. Computer Science Department, University of Rochester, Rochester NY 14627
- Lauesen S. Software Requirements Styles and Techniques. 电子工业出版社, 2002

(上接第 99 页)

该 URP 平台为“1+7+1”(一个门户、七(N)个应用、一个平台。1 个门户: 是指让用户通过一个统一的门户/界面、一次认证, 可以访问整个网络的应用。1 个平台: 是指将所有的应用都集成在一个统一的 URP 平台上^[7]。其中数据中心和门户服务的设计和构建采用了基于视图的方法, 取得了良好的效果。

参考文献

- 殷娜, 傅明. 数字校园的建设发展与规划. 兰州铁道学院学报自然科学版, 2002, 4(8): 142~144

- 蒋东兴, 史宗恺, 等. 大学资源计划的方案研究. 清华大学学报自然科学版, 2004, 4(44): 572~576
- 蒋东兴, 许庆红, 谢聆, 等. 高校信息化建设的一体化思路与实践. 清华大学, 2004
- 黄卫卫, 陈怀楚, 高国柱, 等. 基于交换的数据存储模式研究. 清华大学计算机与信息管理中心, 2004
- 王家顺, 胡耀光, 王田苗, 等. 企业分布式信息共享和集成技术. 微计算机应用, 2001, 7(4): 213~217
- 沈培华. 高校信息化研究与实践. 见: 全国高校信息化研究 2003 年会
- 杨丹, 等. 重庆大学虎溪数字化校园规划方案 1.5 版. 2005