

基于 SOA 的高校档案资源管理系统研究与实现

文/向禹 吴世明

摘要

对当前档案资源管理系统需要解决的主要问题进行分析,从系统、业务、功能、流程等需求分析入手,对系统集成和数据整合等方面进行研究,运用 SOA 架构,采用 MVC 模式开发,研究实现了一种集成化的高校档案资源管理系统,以适应大数据时代的档案工作。

【关键词】SOA MVC 档案管理 系统集成 权限控制

大数据时代的到来,给档案行业带来了机遇和挑战。档案工作的对象不再局限于纸质档案及其电子文件,而是要面向档案资源,档案资源包括档案、电子文件、数据等。工作对

象的变化,促使管理意识和思路、技术手段等都要发生转变,才能适应社会发展,才能推动行业发展。传统的档案管理系统大多是为了解决档案目录管理,工作流程设计简单,已经不能满足当前档案信息化的需求。基于 SOA 的高校档案资源管理系统是集成化、流程化的系统,为解决传统档案管理系统的发展问题而研究设计。

1 系统分析

需求分析是系统研发的基础,系统项目实施的关键。据有关机构分析,软件产品的不正确性、不完整性等问题 80% 以上是需求分析错误所导致,所以需求分析是系统开发的最为关键的一步。

1.1 系统需求

系统到底需要解决哪些主要问题:①系

统集成,通过系统对接,将档案数据产生的业务系统集成起来,以实现档案数据资源的自动归档和整合;②档案数据全文检索,由于档案的特殊性,档案系统应当注重查全率,兼顾查准率,而实现全文检索是提升查全率的重要途径;③工作流程化和规范化,对档案业务流程进行设计,实现流程的自动化,推进档案工作规范化;④权限控制和安全设计,档案资源管理要解决保密与开放,方便使用与用户权限的矛盾,必须做好权限控制与安全设计。

要很好的解决这些问题,须采用相应的技术来实现。SOA 是面向服务的架构,具有服务的透明性、松耦合、可重用、粗粒度、标准化的特点[2],服务独立于系统平台和编程语言,不管其它被集成的业务系统如何变化,档案管理系统提供的集成接口可以不变;双层 PDF 文档技术和 Lucene 规范为解决档案全文检索提供了很好的思路;工作流引擎技术是解决工作流规范化、系统流程化的手段;MVC

<< 上接 80 页

入与产出、信息化建设风险都成为企业信息化建设考量的重要参考因素,而企业云可以降低信息化建设和运营维护成本,降低企业信息基础设施的投资风险。

3 云计算在公共服务方面的应用

各级政府可以将资源整合形成与公众切身利益密切相关的社保云、健康云、教育云等等,教育云能整合开发和引进各类优质教育资源,促进资源普及和共享,解决教育资源不均衡;同时建立基于云计算的图书馆[6]、艺术馆等公益性文化设施,提高文化资源的利用率。健康云可以整合公共卫生、医疗服务等功能的应用,实现不同地区居民能根据云端的居民健康档案和数字化亿元实现卫生自能化体系建立,从根本上解决医疗资源不平衡、看病难、看病贵等问题;就业社保云能解决居民就业和社保全国流动问题。

4 云计算在电子政务方面的应用

政务云能促进跨部门、跨区域、跨层级信息共享和业务协同,实现宏观经济调控、城乡社会管理、应急指挥的快速处置和统一调度。提升服务型政府的服务水平和办事透明度

5 云计算在城乡信息化建设方面的应用

现代农业要充分发挥农村综合信息资源开发利用,推动现代农业的信息化,加快现代农村信息技术应用和科研水平。目前北京市农林科学院农业科技信息研究所已经就云计算在我国农业信息服务中的应用做了一定的研究[7],云计算可以将农业科技信息综合服务提供给涉农用户,将地域行资源转为云端资源,提高解决农业问题的能力,同时也可以将食品质量安全信息放在云端,完成对食品质量全程监控;云计算还可以涉及到农业科研领域。

6 云计算在信息产业方面的应用

云计算对于传统 IT 行业与管理软件行业发生了比较大的变革,对软件业的改变不仅体现在软件的开发方式上,更为重要的改变是软件和服务之间的界限逐渐模糊,特别是随着智能终端的发展,云计算将成为用户从不同智能终端到数据中心的一个良好的沟通,云端即能实现信息系统对不同智能终端需求的响应。

云计算在物流信息资源共享和物流企业与制造企业的业务协同能发挥信息存储、终端互联,促进了以物联网技术为代表的新兴信息技术的发展。

参考文献

[1] 肖连兵,黄林鹏. 网络计算综述,计算机

工程[J],2002,28(3):1-3.

[2] MICHAEL A, ARMANDO F, REAN G, et al. Above the clouds:a Berkeley view of cloud computing [R]. University of California,2009.

[3] 孙在全. 云计算对信息检索的影响及存在的问题,科技情报开发与经济[J].2010,20(31):107-04.

[4] MELL P, GRANCE T. The NIST definition of Cloud Computing[Z] 2011 (Special Publication 800-145).

[5] 魏生辉,郑依华,南凯. 基于云服务的会议服务平台研究与实现,计算机工程[J].2012,38(4):233-238.

[6] 陆颖隽,郑怡萍,邓仲华. 美国图书馆的云服务,图书与情报[J].2012,3:16-21.

[7] 魏清风,罗长寿,孙素芬,张峻峰. 云计算在我国农业信息服务中的研究现状与思考,中国农业科技导报[J].2013,15(4):151-155.

作者单位

1. 湖南女子学院 湖南省长沙市 410004
2. 长沙有色冶金设计研究院有限公司 湖南省长沙市 410011

设计模式也具备高重用性、松耦合、可移植和可维护、低成本高效率等特性 [3]。因此,采用 SOA 架构、MVC 设计开发模式是实现集成化的档案资源管理平台很好的方式。

1.2 高校档案业务分析

档案资源管理是高校管理工作的最后环节,档案及数据的重要性不言而喻,涉及到高校教学、科研、管理、基础建设等工作的方方面面。每一类档案资源的属性各不相同,著录的信息也有差别,因此在设计系统时,根据类别属性不同,设计出不同的视图(即录入界面和浏览界面)。各类文书档案具有相同的属性,因此可采用同一视图;而基建项目档案需要著录项目名称、建设单位、施工单位、设计单位等信息;声像档案则需要著录摄影者、摄影时间、照片人物、地点等信息。

根据业务情况分析,为了系统设计方便,按视图的不同,在系统中将档案划分为文书档案(包括党群、行政、教学、科研、外事、出版、财会、基建、产品、设备)、会计档案、科研档案、声像档案、设备档案、基建项目、学生档案。

1.3 功能需求分析

集成化的档案资源管理系统,必须具备档案资源采集、档案管理、档案服务、在线编研、数据管理、全文检索等子系统,还须具备专业检索、数据安全、统计等功能等。档案资源采集子系统是集成其它部门业务系统的端口,数据采集有手工录入、系统自动采集、档案征集等方式,根据当前信息技术及档案管理发展趋势,考虑接收和采集电子文件及其元数据,并对元数据进行管理;还须具备自动化著

录和标引、自动化组卷;档案管理子系统包括档案移交过程、档案整理、档案鉴定、库房管理等过程。

1.4 业务流程分析

SOA 最重要实现的就是一个完整的业务流程。通过服务快速编排和组合灵活地找寻一些新的业务流程,支撑业务流程的是一组可重用的服务,而服务就是对一个业务功能的封装。档案是特殊的信息资源,业务管理过程中,有大量隐含的业务流程,电子档案处理流程管理非常适合于用工作流管理技术来描述和管理,不仅实现了电子档案处理的自动化流程管理,还可以根据档案归档过程的实际需要灵活地制定电子档案管理流程,适应了电子档案管理非线性运动的特性,大大提高了电子档案归档的规范性。定制业务流程在传统档案软件中是不能实现的,在系统中引入工作流技术,实现档案业务处理流程的可视化、流程自动化、人员操作规范化。

2 SOA服务设计和实现

系统采用基于 SOA 多层框架的技术构架,把相关模块等进行统一的服务包装和整合。基于该架构,系统整体逻辑架构从纵向分为四个层次,基础架构层、企业应用层、企业服务层、信息展现层。各层之间采用 TCP/IP 协议或 Web Service 方式进行数据交互。

2.1 服务设计原则

系统设计开发时,需要先根据业务需求设计出系统服务,借助服务完成模型设计,通过服务间协作与组合形成业务流程,最后完成系统的业务功能。服务设计原则:粗粒度的接

口、无状态服务设计、明确定义的接口。

2.2 数据交换实现

高校各业务系统,大多是根据自身的需求采用不同的开发平台、架构技术、数据结构以及开发语言,在开发之初并没有设计供其它系统集成或者扩展的接口。而档案管理信息系统的业务要求,却要从每个业务系统是获取电子数据和电子文件,以实现纸质文件和电子文件同步归档;实现网络自动化归档,减轻繁杂的档案著录工作。数据交换平台,解决了不同业务系统间数据资源共享的集成方式。如图 1。

采用 SOA 为解决方案,设计粗粒度 Web Service 服务,供所有遵循协议的接口调用,同时设计了线程轮询服务,处理接收到的数据。各业务系统根据协议开发接口,向“数据接收服务”推送数据,数据为 ZIP 包,包含档案管理系统所需要的 XML、txt、PDF 文件。“数据接收服务”只负责接收数据,然后将数据写入中间数据库,然后再由数据文件处理线程对中间数据库轮询,将数据一条条进行解析,写入公用数据库。供档案管理系统进行自动化组卷等操作。

2.3 Web服务接口设计

Web services 是最标准化和最具扩展性的一种 SOA 架构方式,本小节以远程教育学籍管理系统传送数据为例来进行无状态 Web 服务设计。

- ① 接口属性: String[] pushArchiveData (String xmlName,String college,String dataType,byte content)。
- ② 安全性: 支持 WS-Security 规范下: UsernameToken Timestamp Signature Encrypt

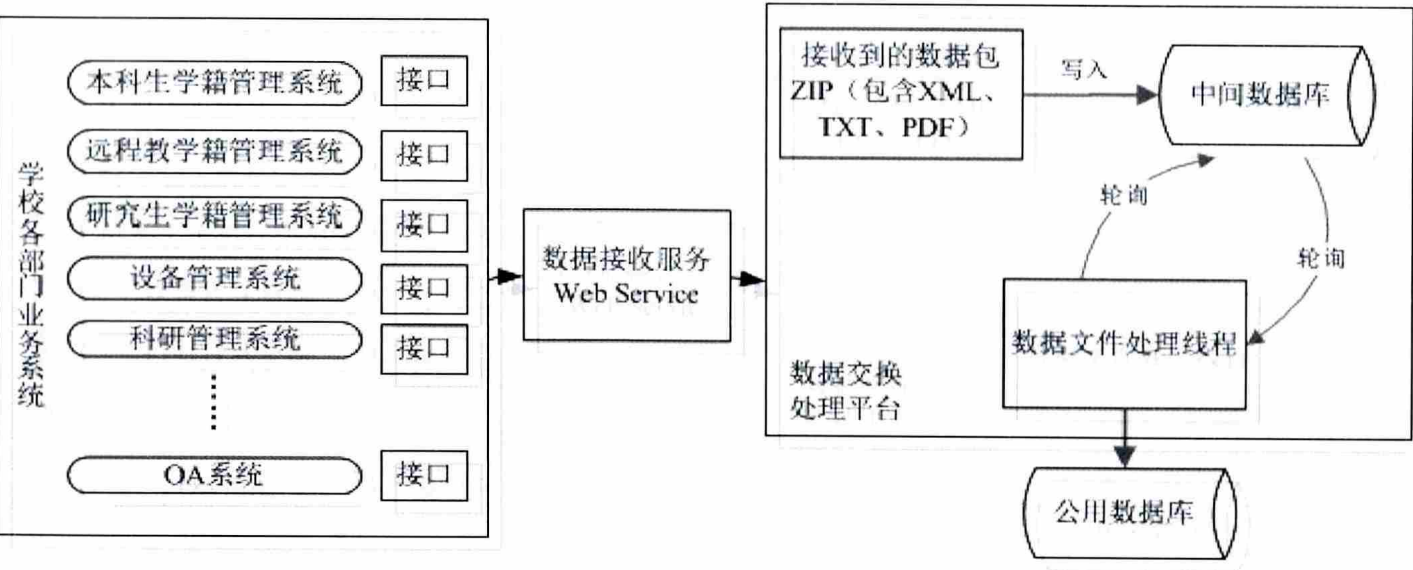


图 1：数据交换示意图

注：默认使用 UsernameToken。

④ 输入参数：xmlName,college,dataType 的类型都为 xs:string；content 类型为 xs:byte。消息传输采用 MTOM 机制，从而使所传输的消息较小。

④ 输出参数：import，类型为 xs:string。返回处理成功的文件标识 id 数组。

⑤ 接口编号对照表：网络教育学院集成业务编号为 wlxy，接口编号 xsxjb(学籍成绩表)、xslqspb(录取名册)、xsbymcb(毕业名册)。其它接口略。

3 MVC开发模式实现

系统采用 MVC (Model-View-Controller) 设计模式开发，MVC 的好处是减弱了业务逻辑接口和数据接口之间的耦合。表现层采用 Struts2 框架；应用层的开发采用 Spring 框架；持久层的开发采用 Ibatis；Web Service 组件服务的开发采用 XFire 框架。子系统、功能模块及整体都基本采用 Struts2+Spring 框架的 MVC 设计模式实现。如“全宗管理”模块的 MVC 设计模型：View 为 Index.jsp，Model 为 Fonds.java，Controller 为 FondsAction.java。

运行过程分析：① index.jsp 页面开始发起请求；②通过 struts2 框架里的 struts.xml 配置文件，查找到对应的 FondsAction.java 控制器；③通过 spring 框架里的 applicationContext-service.xml 配置文件，查找到业务实现 beanFondsManagerImpl.java；④在业务的实现类里，利用 ibatis 框架，通过 FondsMapper.xml 文件实现与数据映射成 javabean(Model)；⑤把对应结果返回到与用户交付的页面。

3.1 MVC开发配置部署

为了使系统能够自动调用 Struts2 框架，首先必须在 web.xml 文件中对 Struts2 进行描述。此外，还需要对 Struts2 框架的配置文件 Struts.xml 进行配置，主要是对框架中的 Action 进行信息描述，使不同的用户请求根据定义来调用相应的 Action，实现页面的跳转。

ActionContextCleanUp 在集成插件方面作用很大，可以使后面的 filter 正确地访问 struts2 的上下文；spring 中提供了一个名为 org.springframework.web.util.IntrospectorCleanUpListener 的监听器。introspection 所使用的缓冲在分析完一个类之后会被马上从 javaBeans Introspector 缓冲中清除掉。监听器的主要作用就是负责处理由 JavaBeans Introspector 的使用而引起的缓冲泄露问题。

3.2 采用JSON数据格式

数据从业务层返回到页面有很多种方式，

本系统用到的是 JSON 方式，所以在 index.jsp 页面中不需要 struts 框架的标签代码。所有数据信息以 JSON 格式和语法，通过变量传递。方便前台数据操作；数据体积小，传输更快。然后再通过 Struts2JsonHelper.java 封装第三方类。

JSON 数据格式为“名称：数值”，全宗模块的某一全宗数据表示如下：

```
{"fondscod":"ZD","fondcname":"中南大学","id":"54","idList":[],"unitaddr":"","unitname":"中南大学","unitperson":"","unittel":""},{ "dataFlag":2,"description":""," }
```

3.3 编写 Action 类和运用 AJAX

编写 Action 类是 MVC 设计的重要一步。而 AJAX 用于提交页面上的数据，异步请求处理，是一种用于创建更好更快以及交互性更强的 Web 应用程序的技术。

4 系统权限控制设计

访问控制是档案管理信息系统重要的部分，由于档案的特殊性，即使全为非涉密档案，但也不宜完全对任何开放；每个角色和用户，甚至每个管理员，每个立卷部门用户对档案数据的不同状态、不同流程、不同字段、不同的功能都有着不同的访问权限。本系统权限控制设计主要是基于角色、权限控制字段的访问控制，基于角色的访问控制具有广泛的适应性和更强的灵活性，角色权限的变化比用户权限的变化要小，授权方便，而且灵活地支持安全策略。

4.1权限控制的粒度问题

根据对系统资源访问能力控制程序的不同，将权限控制划分为精粒度权限控制和细粒度权限控制[6]。粗粒度权限控制：表示类别级，仅考虑对象类别 (Type of Object)，不考虑对象的个体实例，比如角色权限。细粒度权限控制：表示实例级，考虑对象的具体实例 (Instance of object)，比如用户的某个功能按钮。从用户角度来说，粗粒度实现某一用户组 (角色) 的共同权限，细粒度实现具体用户权限。

4.2用户、角色、权限的关系

用户和权限的关系为多对多的关系，一个用户可以属于多个角色，一个角色也可以包含多个用户。角色和权限仍然是多对多的关系，一个角色可以用多个权限，一个权限可以属于多个角色。

4.3权限控制字段

角色权限控制的粒度较粗，如果角色划分过细又会使角色泛滥。本系统做了一项重要设计，对字段增加了“权限控制字段”属性，

在设计或者修改数据表结构时，勾选即为权限控制字段。保管期限、密级、部门、全宗、分类号、输入人等都可以设置为权限控制字段。

4.4用户权限设计

系统权限分为功能权限、数据权限、流程权限。同一类权限之间为逻辑 AND 关系，如部门和全宗两个字段的修改权限都赋予 A 角色，那么，在数据满足“部门 and 全宗”条件下，才能对数据修改。不同类权限之间，角色与用户之间的权限为逻辑 OR 的关系。系统用户权限管理功能设计了 11 个数据表：t_function 功能权限表，t_role 角色表，t_roledata_func_action 角色数据功能表，t_user 用户表，t_userdata_func_action 用户数据功能表，t_userfunction 用户与功能映射表，分级授权表，用户查看流程权限表，角色查看流程权限表。

系统总体遵循 J2EE 规范，主要采用 SOA 服务架构和 MVC 设计模式来开发系统，同时运用了多层框架和技术，使系统具备先进性、实用性、可扩充可维护性、可靠和安全保密性，也使系统具备了可持续发展的能力。研究基于 SOA 的高校档案资源管理系统对于大数据时代的高校档案资源管理具有重要的现实意义。

参考文献

- [1] linFen. 软件需求分析方法 [EB/OL]. <http://www.cnblogs.com>, 2009-06-18
- [2] 金竹青. 基于 SOA 架构的 B2B 电子商务平台的设计与研究 [D]. 大连：大连海事大学，2008.
- [3] mvc 原理和 mvc 模式的优缺点 [EB/OL]. ITeye 网博客 [2012-07-7].
- [4] 向禹. 基于 SOA 的高校档案管理信息系统设计与实现 [D]. 长沙：中南大学，2013.
- [5] 赵莉娟. 基于 SOA 架构的高校数字档案信息资源整合的研究与设计 [D]. 武汉：武汉理工大学，2012.
- [6] 湛勇，许大丹. 基于角色的粗粒度权限控制框架设计实现 [J]. 华北水利水电学院学报，2007(1):66-68.

作者简介

向禹 (1976-)，男，副研究员，硕士。中南大学档案馆信息技术部主任。研究方向为数据库技术、信息处理技术、档案信息化。吴世明 (1969-)，男，副研究员，硕士，中南大学档案馆副馆长。

作者单位

中南大学档案馆 湖南省长沙市 410083