

EAST 文档管理系统的设计与实现

黄 静,王 枫,谢志新,肖炳甲

(中国科学院 等离子体物理研究所,安徽 合肥 230031)

摘 要:为了更好地管理 EAST 的项目文档和实验文档,以促进实验的顺利运行及各部门之间的相互协作,专门研究与设计了一个适用于 EAST 的基于 B/S 结构的文档管理系统。系统采用 UID 作为文档的唯一标识,通过文档 UID 可以直接检索定位文档。Web 前端运用 Ajax 技术与 jQuery EasyUI,实现页面的局部更新。结合开源 PHP 框架完成了系统的设计开发,最终实现了文档分类存储、在线操作文档、文档版本控制、文档权限管理及工作流等功能。该系统已成功应用于 EAST 实验中,极大地提高了各部门的工作效率。

关键词:文档管理;B/S 结构;文档 UID;分类存储;版本控制

中图分类号:TP302.1

文献标识码:A

文章编号:1673-629X(2014)07-0013-04

doi:10.3969/j.issn.1673-629X.2014.07.004

Design and Implementation of EAST Document Management System

HUANG Jing, WANG Feng, XIE Zhi-xin, XIAO Bing-jia

(Institute of Plasma Physics, Chinese Academy of Sciences, Hefei 230031, China)

Abstract: In order to manage the project and experimental documents of EAST tokamak, so as to promote the project progress and the collaboration between departments, a special electronic document management system based on browser/server structure is designed. All documents can be retrieved and located through document UID which is the unique identification of a document. The Ajax technology and jQuery EasyUI framework are adopted to realize partial page refresh. The system is developed based on CodeIgniter which is an open-source PHP framework, and main functions of the system have been implemented including document classification storage, on-line operation, version control, user and permission, working-flow management, and other functions. This system has successfully been applied in EAST experiments, greatly improving the work efficiency in each department.

Key words: document management; B/S structure; document UID; classification storage; version control

0 引 言

信息化为人们的生活带来了极大的方便与快捷,同时也能强烈地感受到信息化时代带来的巨大压力和挑战,信息流通速度大幅度提高,文档数量也迅速增加^[1]。由于计算机和网络的普及,几乎每个人都会或多或少地参与文档管理工作,成为文档管理系统的用户。EAST (Experimental Advanced Superconducting Tokamak) 是世界上第一个非圆截面全超导偏滤器托卡马克,是用于研究核聚变的一个重要科学实验装置,可开展受控核聚变相关前沿物理问题的研究^[2]。随着 EAST 实验的运行和项目的进展,有大量的实验文档、项目文档、会议文档等需要进行管理。在管理文档的过程中,可能会遇到海量文档存储困难、检索效率低下、文档无法实时共享、文档安全性得不到保障等一系

列问题。针对上述问题,需设计与实现一个适用于管理 EAST 相关文档的文档管理系统。EDM 是 EAST Document Management 的简称,是用来管理与 EAST 相关文档的一个在线文档管理系统。为了保证系统能够满足日常工作需要,有效地提高文档管理水平,功能上须达到以下目标:文档集中安全地存储,文档在线操作,文档版本控制以及文档权限管理^[3]。而非功能性需求则包括系统的易用性、稳定性、可扩展性^[4],以方便后期系统的维护及功能的完善。在这样的需求分析下,完成对 EDM 这一系统的设计及实现。

1 EDM 总体架构

EDM 采用由表现层、逻辑层和数据层构成的三层 B/S 结构,分别是客户端、Web 服务器和数据库^[5]。基

收稿日期:2013-10-07

修回日期:2014-01-12

网络出版时间:2014-04-24

基金项目:国家“973”重点基础研究发展计划项目(2012GB105000,2011GB101000)

作者简介:黄 静(1988-),女,硕士研究生,研究方向为 Web 应用、数据挖掘;王 枫,博士,高级工程师,研究方向为数据采集、数据服务等。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140424.0839.084.html>

于浏览器的客户端负责与用户进行交互,处理极少部分的事务逻辑,主要事务逻辑通过浏览器发送请求交给 Web 服务器处理;Web 服务器接收请求,通过多种方式与数据库(MySQL)进行数据交互,最后把处理结果返回给客户端;数据库主要用来存取数据,处理 Web 服务器的数据请求并返回给 Web 服务器。系统的总体结构如图 1 所示。

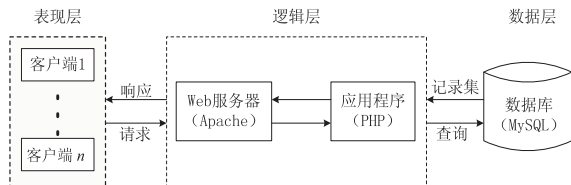


图 1 系统架构图

系统功能主要分为两大模块:文档管理和用户管理。文档管理主要负责文档信息管理及文档在线操作等功能。用户管理包括统一用户信息管理,文档权限分配及 workflow 等功能。统一用户管理主要是对用户信息进行集中的添加、修改、删除等操作,文档权限分配是指不同用户角色拥有对文档或目录的不同操作权限,workflow 是用来管理文档的流程,包括文档的创建、评阅、审批、发布^[6]。文档管理是系统的核心功能,主要有文档分类存储、文档浏览、文档上传下载、文档检索、文档版本控制、文档的权限及 workflow 等功能。EDM 功能结构如图 2 所示。

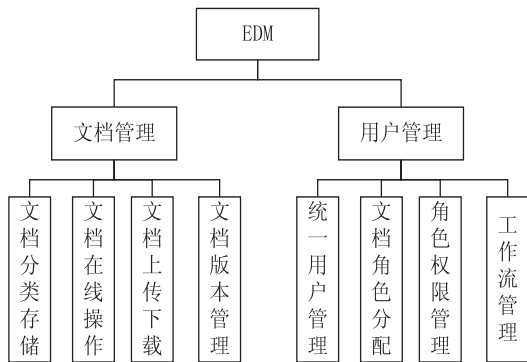


图 2 系统功能结构图

基于 EDM 的功能需求,使用 Eclipse 的插件 ER Master 建立 MySQL 数据表,并根据各个表的关系创建了 ER(Entity Relationship)模型图。EDM 前端引用 EasyUI 插件,它是基于 jQuery 的用户界面插件的集合。后端应用了开源框架 CI(CodeIgniter 的简称),它使用 MVC(Model, View, Controllers)的设计模式,是一个基于 PHP 的轻量级框架^[7]。

2 EDM 关键技术

2.1 文档 UID

UID 是 EDM 中所有目录和文档的唯一标识,定为 6 位,由 26 个英文大写字母和 0~9 数字组成,如 ABC-

DEF。它可以提高系统效率,并且容易识别与记忆。通过 UID 可以直接定位目录或文档所在位置,以便检索、查阅等操作。有多种方法生成 UID,主要分为预先生成和在线生成两种。

(1)预先生成就是按照一定的算法生成 UID,即在建立数据库时就创建所有的 UID,并分开存储已使用的 UID 和未使用的 UID。共涉及到 uid_available、uid_used 和 uid_reserved 三个表,uid_available 用来存储可申请的 UID,uid_used 存储已使用的 UID,uid_reserved 存储已申请但未使用的 UID。

(2)在线生成就是在创建文档或目录时,按照相应的算法在线生成 UID,生成的 UID 不能与已用过的 UID 重复,否则重新生成。为了增大在线生成 UID 的随机性,按一定的算法取随机整数值,再通过进制转换把十进制转为三十六进制,最终生成形如 ABCDEF 的 UID。用表 uid_used 来存储已使用的 UID,uid_reserved 保存用户已申请但未使用的 UID,只为每个用户保存一个已申请的 UID。

作为唯一标识的身份证号、手机号、车牌号等都是采用预先生成的方法,而随机验证码的产生是采用在线生成的方法。由于预先生成 UID 的方法涉及到三个 UID 数据表,使得 UID 的管理也比较复杂,最终系统采用在线生成的方法来获取 UID,获取 UID 的流程如图 3 所示。

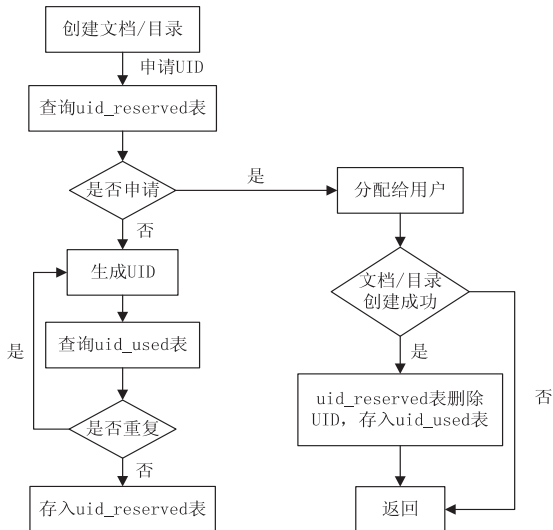


图 3 获取 UID 流程图

随着文档数量的增加,数据表中可能会有上百万条甚至更多的记录,此时需要用到 MySQL 的查询优化技术,以便迅速存取及检索 UID。

2.2 局部页面刷新技术

在浏览 Web 页面的时候,很多情况下要保留页面的公共部分,只刷新部分页面内容^[8]。Frameset(框架)是一种比较传统的、易于掌握的页面刷新技术,它可以将 HTML 分为多个页面,即一个 Web 页面可由多

个框架页的组合来显示。重载页面时只需要重载页面中的一个框架,而不需要重新加载整个页面。因此会产生很多页面,不便于管理,多框架的页面也会增加服务器的负担,影响页面加载速度。另一种技术 Ajax (Asynchronous JavaScript and XML) 采用异步交互方式,按需求与后台服务器进行少量的数据交换,不影响用户操作,实现页面的异步更新^[9]。EDM 的界面设计应用了 EasyUI 框架,只需通过 HTML 标签来定义用户界面,不用写很多复杂的 JavaScript 代码,可以简单快速地实现界面布局,并统一界面风格。但是 Ajax 返回的页面无法自动渲染 EasyUI 的样式,造成界面风格不一致。解决方法是:把需要刷新的部分定义成一个 panel,通过 refresh 方法加载不同 URL 返回页面,实现页面的刷新。

具体实现:
HTML 页面中定义: id = “*” class = “easyui-panel”
JS 中返回: \$(‘#id’).panel(‘refresh’, ‘url’)
EDM 采用 Ajax 同时结合 EasyUI 来实现局部页面的刷新。如不同用户或不同目录对应不同的权限菜单是通过 Ajax 实现的,而菜单对应的具体操作则是通过 EasyUI 中的 panel 实现的。

3 系统功能实现

3.1 文档分类存储

EDM 通过系统直接存储文件和相关元数据,根据文档的类别、格式等不同,对文档分类进行存储,方便文档随时随地获取,促进相互之间的交流与合作。EDM 分为 Project、Meeting、Department、Test 四个类别,其中 Meeting 主要用来存放会议报告、会议记录等文档,Project 按项目不同来存放与项目相关的文档,Department 是根据部门不同来存放用户的个人工作文档,Test 是用来测试的目录。EDM 采用树形目录结构来存储文档和目录,但是 MySQL 关系数据库是以二维表的形式记录存储数据信息,所以不能直接将树结构存入数据库,需要设计合适的架构及相应的 CRUD (Create, Read, Update, Delete) 算法来存储树形结构。表结构设计为 {id, pid, uid, name, path}, 这样直接记录了节点之间的继承关系,非常直观和方便,如表 1 所示。

- (1) Create: 在当前节点增加新节点时, id 是在线生成的, uid 是 id 通过三十六进制转换得到的。pid 为当前节点的 id, name 为新节点的名字, path 则是当前节点的 path+新节点的 id。
- (2) Read: 点击当前节点可以获取一级子节点信息,并能获取当前节点及其父节点的信息,通过 path

可以得到其在树结构中的路径。

表 1 目录结构节点关系表

id	pid	uid	name	path
1	0	000001	EDM	0
2	1	000002	Meeting	1,2
3	1	000003	Project	1,2
4	2	000004	Project1	1,2,4
5	2	000005	Project2	1,2,5
6	1	000006	Department	1,2
7	1	000007	Test	1,2

- (3) Update: 可以修改节点的名字及父节点,父节点修改后,路径及其所有子节点的信息都要随之改变。
- (4) Delete: 删除某一个节点,需要递归删除所有子节点再删除当前节点。
- 树形目录结构中每一个目录都是其上一级目录的子目录,同时是下一级目录及文档的父目录,直观地记录了各个节点之间的关系。在 EDM 中,目录和文档都是作为节点存储在树形结构中,这样对每一个文档的定位,都可以从根目录开始,由一层层目录的节点编号构成的路径来实现。通过树形结构,对相关的文档进行分类存储,方便文档的管理,及更加快速地查找定位文档。EDM 支持目录扩展,每个子目录也可以根据需要再进行无限制地划分。

3.2 文档在线操作

EDM 支持用户在线浏览文档、创建文档、编辑文档、删除文档、上传下载文档等基本的文档管理操作。系统支持 word、txt、pdf、jpeg 等常见文档格式的在线查看。创建一个文档时,系统会自动分配一个 UID,用户可以在线编辑文档标题、类型、作者、创建日期等基本描述信息及文档内容,还可以上传附件供浏览下载。删除一个文档会删除所有的元数据及其附件,具有相应权限的用户可以执行这一操作。

文档的检索是针对所有已经发布的文档,未发布的文档是不能检索到的。系统支持的检索方式可分为按文档属性检索和按文档内容检索两种。

- (1) 按文档属性检索是指根据文档的 UID、类型、作者、发布时间等来检索文档,直接与数据库中存储的数据进行匹配得到检索结果。
- (2) 按文档内容检索采用全文检索技术^[10],对文档的关键词建立索引,根据用户输入的关键词进行分析匹配,检索出相关文档^[11]。

3.3 文档版本控制

一个文档可以有多个版本,每次修改完成之后都作为一个新的版本保存下来。对于每个文档,默认查看的都是它的最新版本,即当前版本,在查看当前版本文档的元信息时,如果有历史版本则会看到相关历史

版本的信息,通过不同的版本号来查看历史版本,并可下载到本地,方便对各个版本间的差异进行比较^[12]。EDM 中,对文档各个版本的管理如图 4 所示的流程。

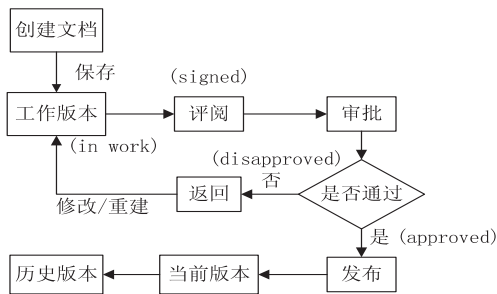


图 4 版本控制流程图

在完成一个文档的创建工作之后,此文档作为工作版本保存下来,并提交给具有相关权限的人员进行审批,直到审批成功,才能进行发布,发布的文档作为文档的当前版本供查阅、编辑等操作^[13]。每次的修改都是针对文档的当前版本进行,修改后的文档也同样需要经过工作版本-审批-发布-当前版本的流程。文档的状态(分为四种:in work, signed, approved, disapproved)会随着流程的转变而变化,系统集成了电子邮件的功能,会以邮件通知相关评审、审批人员,同时也会通知作者关于文档的状态改变。修改并发布成功后,文档版本号自动改变(例如 1.1 变为 1.2)。文档的版本管理涉及到版本控制技术和工作流技术,是文档管理中不可或缺的一项重要功能。

3.4 用户权限管理

EDM 采用 LDAP(Lightweight Directory Access Protocol)对用户进行统一管理。用户登录系统时,需要进行身份认证,根据用户身份自动分配权限。LDAP 作为 EDM 的认证数据库,存储用户的基本信息,包括 id, name, password, E-mail 等。用户通过浏览器输入用户名和密码,从 LDAP 中读取用户的认证信息,认证通过后,用户即可进入系统并获得相应权限^[14]。

对于不同目录或文档,不同的用户所拥有的权限不同。系统通过采用基于角色的访问控制技术(Role Based Access Control, RBAC)^[15],来实现文档的权限管理。目录和文档对应不同的角色信息,目录的角色包括目录管理员(Admin)、创建者(Creator)、浏览者(Viewer);文档的角色包括作者(Author)、评阅人(Reviewer)、审批人(Approver)。每个角色对应一种操作权限,不同的用户或用户组根据其角色的不同对文档具有不同的操作权限。

4 结束语

文中针对文档管理系统的需求分析,结合使用 CI 框架及相关 Web 技术,完成了系统的设计与开发。

EDM 还实现了文档的存储、查阅、检索等基本功能,不同的用户角色对文档具有不同的操作权限,以及文档的版本控制和工作流功能。EDM 已成功地应用到 EAST 实验中,极大地提高了管理 EAST 相关文档的工作效率。系统的各子功能模块是相互独立的,有良好的扩展性,可根据需求不同来增加新功能。EDM 将会不断加入更多功能,如支持多语言,在线文档模板等。

参考文献:

- [1] 赵宇. 如何推进企业文档管理工作的思考[J]. 中国商界, 2009(5): 231-232.
- [2] Song Y T, Yao D A, Wu S, et al. Structural analysis and manufacture for the vacuum vessel of experimental advanced superconducting tokamak (EAST) device[J]. Fusion Engineering and Design, 2006, 81(8): 1117-1122.
- [3] 刘丽, 吴秋云, 李军. 基于 Web 的分布式文档管理系统的设计与实现[J]. 计算机工程与科学, 2007, 29(1): 14-16.
- [4] 蒋琨, 季振山, 肖炳甲. 基于开源技术的 Web 协作工作软件设计实现[J]. 微计算机信息, 2012, 28(10): 406-408.
- [5] 叶宝军, 秦鹏飞. 文档管理系统模型研究与应用[J]. 东华大学学报(自然科学版), 2002, 28(5): 60-63.
- [6] 刘伟, 朱一凡, 魏洪涛. 工作流技术在办公自动化中的应用[J]. 计算机工程与设计, 2006, 27(7): 1308-1310.
- [7] 杨彦侃, 谭心. 一种基于 CodeIgniter 框架科研论文管理系统的设计与实现[J]. 计算机应用与软件, 2012, 29(12): 223-224.
- [8] 卢成均. 通用集中式文档管理系统的设计与实现[J]. 计算机工程与设计, 2006, 27(24): 4792-4796.
- [9] 徐驰, 徐燕凌. Ajax 模式在异步交互 Web 环境中的应用[J]. 计算机技术与发展, 2006, 16(11): 228-230.
- [10] Ren Zujie, Chen Ke, Shou Lidan, et al. HAPS: supporting effective and efficient full-text P2P search with peer dynamics[J]. Journal of Computer Science & Technology, 2010, 25(3): 482-498.
- [11] 方艳芬. 多源文档全文检索系统设计与实现[D]. 武汉: 华中师范大学, 2009.
- [12] Zhong Shisheng, Li Tao. A version management model of PDM system and its realization[J]. Journal of Harbin Institute of Technology, 2008, 15(4): 459-464.
- [13] 王红丽. 协同设计中版本控制技术的研究[D]. 哈尔滨: 哈尔滨工程大学, 2006.
- [14] 贺玉明, 李晋宏, 唐辉. LDAP 在数字校园统一身份认证系统中的应用[J]. 计算机技术与发展, 2011, 21(5): 139-142.
- [15] Zhu Yiqun, Li Jianhua, Zhang Quanhai. General attribute based RBAC model for web services[J]. Wuhan University Journal of Natural Sciences, 2008, 13(1): 81-86.

作者：[黄静](#), [王枫](#), [谢志新](#), [肖炳甲](#), [HUANG Jing](#), [WANG Feng](#), [XIE Zhi-xin](#), [XIAO Bing-jia](#)
作者单位：[中国科学院 等离子体物理研究所, 安徽 合肥, 230031](#)
刊名：[计算机技术与发展](#) 
英文刊名：[Computer Technology and Development](#)
年, 卷(期): 2014(7)

参考文献(15条)

1. [赵宇](#) [如何推进企业文档管理工作的思考](#) 2009(05)
2. [Song Y T;Yao D A;Wu S](#) [Structural analysis and manu-facture for the vacuum vessel of experimental advanced super-conducting tokamak \(EAST\) device](#) 2006(08)
3. [刘丽](#); [吴秋云](#); [李军](#) [基于Web的分布式文档管理系统的设计与实现](#) 2007(01)
4. [蒋琨](#); [季振山](#); [肖炳甲](#) [基于开源技术的Web 协作工作软件设计实现](#) 2012(10)
5. [叶宝军](#); [秦鹏飞](#) [文档管理系统模型研究与应用](#) 2002(05)
6. [刘伟](#); [朱一凡](#); [魏洪涛](#) [workflow技术在办公自动化中的应用](#) 2006(07)
7. [杨彦侃](#); [谭心](#) [一种基于CodeIgniter框架科研论文管理系统的研究与实现](#) 2012(12)
8. [卢成均](#) [通用集中式文档管理系统的设计与实现](#) 2006(24)
9. [徐驰](#); [徐燕凌](#) [Ajax模式在异步交互Web环境中的应用](#) 2006(11)
10. [Ren Zujie;Chen Ke;Shou Lidan](#) [HAPS:supporting ef-fective and effcient full-text P2P search with peer dynamics](#) 2010(03)
11. [方艳芬](#) [多源文档全文检索系统设计 with 实现](#) 2009
12. [Zhong Shisheng;Li Tao](#) [A version management model of PDM system and its realization](#) 2008(04)
13. [王红丽](#) [协同设计中版本控制技术的研究](#) 2006
14. [贺玉明](#); [李晋宏](#); [唐辉](#) [LDAP在数字校园统一身份认证系统中的应用](#) 2011(05)
15. [Zhu Yiqun;Li Jianhua;Zhang Quanhai](#) [General attribute ba-sed RBAC model for web services](#) 2008(01)

引用本文格式：[黄静](#). [王枫](#). [谢志新](#). [肖炳甲](#). [HUANG Jing](#). [WANG Feng](#). [XIE Zhi-xin](#). [XIAO Bing-jia](#) [EAST文档管理系统的设计与实现](#) [期刊论文]-[计算机技术与发展](#) 2014(7)