Short Report on Progress of ETD-db 2.0 Rewrite

Monday, April 25, 2011

Written by Sung Hee Park and Kimberli D. Weeks

1. Introduction

   a. Aims/Objectives
   This paper describes a new version of Virginia Tech’s ETD-db digital library system, ETD-db 2.0, which is a web application with the web pages generated by Ruby on Rails scripts. Objectives of rewriting ETD-db are 1) to improve the original powerful functionalities of current ETD-db, and 2) to provide new reliable and secure features to handle ETD collections. It continues to work with any database (e.g., MySQL, PostgreSQL, Oracle, etc.) and web server host (e.g., Apache, Tomcat).

   b. Methods
   We are using Ruby on Rails, a state of the art agile web development paradigm, as our primary web application development framework for improved maintenance. To strengthen the original functionalities of ETD-db, existing functionalities of the current ETD-db system have been articulated and additional needs also collected to be improved in the requirement specification process. To provide new reliable and secure features, a user group consisting of system administrators, managers, and authors has been interviewed to describe their maintenance/usage experiences. Collected requirements have been analyzed by the use case based requirement engineering approach. According to the functional/non-functional requirements drawn through use case analysis, ETD-db 2.0 prototype has been designed and implemented rapidly and then a test driven development (TDD) approach to ensure safer/reliable codes have been used.

   c. Result:
   As a result, we have drawn additional safer and more reliable features for a new ETD-db system (See http://scholar.lib.vt.edu/theses/NDLTD/BoD201012/ETDdbVer2Features.pdf). The ETD-db 2.0 system prototype will be released in near future. It has integrated a directory system by using a secured LDAP protocol (e.g., OpenLDAP, OpenSSL) to authenticate users. In addition, integrating the ETD-db 2.0 with the Banner (HR) system is under development to enhance usability. We have chosen more maintainable systems administrator functions (e.g., several statistics like usage/status report, logging and audit). Managers have a better approval and release notification method. For instance, on approval of ‘unrestricted’ ETDs, emails are automatically generated for authors, committee members, UMI, when appropriate. At the time of automatic release of a ‘withheld’ or ‘restricted’ ETD, an email is sent to UMI when appropriate.

   d. Conclusions:
   ETD-db 2.0 is designed and rewritten to improve the current ETD-db system in terms of reliability and security giving many benefits to all stakeholders (e.g., library, graduate school, and contributors) and users (e.g., system administrators, managers, reviewers, catalogers, authors, etc.). For our development process, a state of the art web development framework, Ruby on Rails, has been used to support requirements clarified by requirement analysis. Security issues have been improved by fine-grained access control, increased audit logging, and maintenance eliminating inconsistencies between the file structure and database. In addition, more reliable
content management has been accomplished through consistency between contents and their metadata, content integrity, and version control of each ETD. A future version will make the new system more stable by means of repetitive tests, debugging and trouble shooting.

2. Accomplishments and Progresses
   a. Requirements & Documentations
      i. ETD-db 1.0
         1) One single password per role

         In ETD-db 1.0, one single password per role exists even though more than two users have the same role. For a more reliable and safer system, the role management functionalities by admin are required.

      ii. Fine-grained access control

         Currently, ETD-db provides different access permission to different databases. For example, ETD-db 1.0 has submitted ETD database, available ETD database and withheld ETD database. System administrators, managers, or reviewers can access to all ETD databases while authors have access permission to the only submitted ETD database. Different roles should have not just database-level access permission but also digital object-level and action-level access permission for more secure access control.

   b. Design & Documentations
      i. Authentication

         Authentication is a process to check if a user is a registered user. This process will be verified by the centralized authentication web service (e.g., LDAP via ED-Auth for the first release) or a user table in the database which ETD-db 2.0 hosts. If centralized authentication web services are used for authentication, the specific authentication method will verify the user. If the user is successfully authenticated, the process will be passed to the authorization process. If not, the control will be redirected to the ‘Forgot Password’ and ‘Help’ page. This modular authentication design should allow for future integration with other centralized authentication web services (e.g., Google ID, Open ID Shiboleth, etc.).
ii. Authorization

Authorization is a process to verify if a user has an appropriate privilege or not. In this rewriting ETD-db, we have focused on the multiple roles with the same person and different users with the same role. ETD-db 2.0 is designed in order for different users to have the same role (e.g., student workers with reviewer in the graduate school) and login with their own user name and password. In addition, ETD-db 2.0 aims to authorize a user with multiple roles (e.g., admin, manager or reviewer). Figure 2 shows a flowchart for authorization in case of multiple roles.
iii. Submission

Strictly speaking, submission process is different from a login process as an author. Once an authentication as a valid user and authorization as an author role has succeeded, the user can see a ‘new ETD submission’ link and incomplete (pending) submitted ETDs. Presently, to facilitate an author to submit a new ETD, integrating ETD-db with Banner system is designed as a new function for Authors. Basic Information related to a user (e.g., department, degree, document type, etc.) will be able to be provided by an administrative Banner system. Finally, the author can enter his/her information through web interface.
1) Interaction between Objects In ‘show ETDs by author’ View

We are implementing each view (e.g., show, new, edit, delete, etc.) for a specific controller (e.g., submit, review, manage, catalog, admin, search, browse). For example, Figure 4 illustrates a class diagram for ‘show ETDs by author’ view.
c. Implementations
   i. Role management
      1) Admin functionalities
         a. Register Digital Objects
         b. Register Actions
         c. Register Roles
         d. Assignment of Permissions to Each Role
      2) Authorization for Multiple Roles
         a. Different users with the same role
         b. Same user with the multiple roles (Completed)
   ii. Submission process
      1) ETD metadata submission
         a. Getting information from ED-ID (in discussion about access permission)
         b. Getting information from Banner (in discussion about access permission)
         c. Getting information from VT-Specific Banner (in discussion about access permission)
         d. Getting information from Users (completed)
      2) ETD file submission (Completed)
      3) Committee member submission (Completed)

3. Plans
   a. Getting permissions to access the VT Banner system for integration- in discussion
   b. Interviewing with reviewers (Graduate School) to get requirements - in scheduling
   c. Audit Logging, Provenance for seamless communication between author and reviewer – in designing and to be implemented in May and June.
   d. Import & Export Function to be implemented in July.

4. Scenarios
   a. Submission process
      i. Authentication
Log in

Enter your Virginia Tech PID (username) and password. Note that you must use the PID you were originally assigned, and not any aliases you may have created. Help with PIDs (including forgotten passwords or deactivated accounts).

Name
shpark
Password

Log in

Your browser must accept cookies to continue the submission process. (browser compatibility issues).

Please note: If you are using Internet Explorer, and the clock on your machine is more than 15 minutes out of sync with the clock on our server, you may have problems with the cookies used by the submission software. The current date and time on our server is Tue Jul 13 10:03:33 EDT 2010. Be sure to reload this page before attempting to reset your clock.

Figure 5. Login Page for Authentication

➢ Authentication failure

Invalid Authentication

Please refer to Help with PIDs. Please try log in again.

Development Contact Us
URL: https://auth

Figure 6. Authentication Failure Page

Type in a wrong username and a password (show a “Help” link).

➢ Multiple role case
Type in a correct username and a password.

- Author role

Show incomplete ETDs.
Add new ETD.

Admin role
Reenter credentials for authorization.
b. Role management (Admin functionalities)
   - Add roles

   ![Add new role Page](image)
   
   Figure 12. 'Add new role' Page
   
   - Add users
Enter a new person

Role
- Manager
- Admin
- Reviewer
- Cataloger
- Author
- Co-chair

First name

Middle name

Last name

Email

Figure 13. 'Add new user' Page

- Edit Permission per role
# Add new permission

<table>
<thead>
<tr>
<th>Digital Object</th>
<th>Actions</th>
<th>Add</th>
<th>Modify</th>
<th>Delete</th>
<th>View</th>
<th>Approve</th>
<th>Catalog</th>
<th>Import</th>
<th>Export</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etd metadata</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Etd file</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Etd copyright</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etd availability</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etd status</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admin</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reviewer</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cataloger</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 14. 'Add new permission' Page**