SimplyCT Workflow System

Ridhô Jeftha
University of Cape Town
Rondebosch
Cape Town, South Africa
ridho.jeftha@agofdn.net

1. ABSTRACT

SimplyCT is framework for simple online or offline digital archives with preservation, standardisation and extensibility being major foci. A workflow system was developed for the SimplyCT framework to test its applicability as an institutional repository. The workflow system was built as a Web application using Java Servlets and HTML. The workflow system shows potential to assist with the input and manipulation of data.

2. INTRODUCTION

Submission of a file into the SimplyCT repository is currently done manually as SimplyCT was mainly designed for cultural heritage. This tends to be a tedious way of submission as a technical professional and a curator (for approval of the submission) must be physically present at the same time and place. The aim therefore was to test SimplyCT's applicability to Institutional Repositories by adding a workflow management system.

The SimplyCT framework is based on a flat file store rather than a database. There is thus no formal query language (such as SQL) to add data into the repository.

3. BACKGROUND

There are a few existing digital repository systems, such as DSpace[1], EPrints[2] and Fedora[3].

The DSpace system has 8 states (3 being pool states where the file waits to enter a new state) in handling a submission. The item is held in a workspace until it is submitted by a user. All the metadata is then added and the item is held in temporary storage until it is approved or rejected by a reviewer. When a file is approved, it is then entered into the repository, and when it is rejected the submitter receives an e-mail with information regarding the rejection.

DSpace tracks the state of the submission by adding a provenance message to the Dublin Core Metadata that includes the filenames and checksums of the content that was submitted. After each state change, a similar provenance message is added. [4]

EPrints follows a similar design by storing the file in a staging area before it is accepted into the repository by a reviewer.

4. DESIGN OF SOLUTION

The system is designed for two different users. The one user will have access to upload functionality while another user (administrator/curator, etc.) will have privileges to add documents to the repository.

Login System:

The login system has to grant access to different users. As this is an early prototype that is only testing the framework's applicability to Institutional Repositories, no formal user management system has been developed.

A mock login page is used to distinguish between a normal user and an administrator.

File Submissions:

A normal user can submit a file for approval. The Java servlet fetches the file and then constructs the webpage for entering metadata. After the metadata is captured, it is stored in an XML file with the same name as the document submission but with a [.medata] extension.

The file is stored in a staging area and does not enter the repository until it is approved by the administrator.

Administrator:

The administrator can see a list of all the files that are currently in the staging area. A file's metadata can be viewed and edited. The file can also be accepted into or rejected from the repository.

5. IMPLEMENTATION DETAIL

The system was tested using Eclipse Java EE IDE for Web Developers and Apache TOMCAT.

The Login Page is accessed using the server created by Tomcat at `/SimplyCTWorkflowsystem/LoginPage.html`.

The `UploadServlet` manages the uploading of the file and `MetaServlet` manages it's metadata file.

In the Administrator mode the `DisplayServlet` is used to control the actions of the administrator.

6. RESULTS

The seven (7) users who evaluated the system all had prior experience with existing repository systems such as Eprints and Fedora.

Users where asked to upload a file to the staging area as well as to edit files as an administrator. A questionnaire (see appendix A) was given to each user to complete to capture the results.

All of the seven users agreed that the system performs the basic tasks of managing file entry. The response was mostly positive. More users thought that the navigation between tasks was easy and intuitive and that the system offers similar functionality to other repository systems.

7. CONCLUSIONS

The Workflow Management system was found to be effective with the SimplyCT framework.

8. FUTURE WORK

As this was a very small research project a user management system was not implemented in the workflow system. If a secure user management system is integrated into this system, more functionality can be added such as a message that can be attached to files that are rejected or users can track the progress of their submissions.

More functionality can also be added for file submissions, such as different metadata tags for different types of files.

9. REFERENCES

[1] DSpace,
Available http://www.dspace.org/

[2] EPrints,
Available http://www.eprints.org/

[3] Fedora,
Available http://fedora-commons.org/

[4] 2010, *Chapter 13. DSpace System Documentation: Business Logic Layer,*
Available
http://www.dspace.org/1_6_2Documentation/ch13.html#docbook-business.html-workflow

SimplyCT IR Workflow System

SimplyCT is a repository designed for mainly cultural heritage submissions. The SimplyCT framework is based on a flat file store rather than a database. The aim of this project is to establish if SimplyCT is applicable to an Institutional Repository by adding a workflow system on top of it. This is a very early prototype of the system and therefore cannot be compared to DSpace or EPrints.

Please run through a normal submission of a file and answer the questions below:

- Log in using the username: user , and no password
- Choose a file for submission
- Enter metadata for the file submission
- Log out or add another file

- Log in using the username: admin, and no password
- View one of the submissions
- Edit one of the submissions' metadata and view if the change has been made
- Reject a file from entering the repository (check in actual folder)
- Accept a file from entering the repository

SimplyCTEvaluation:

With which Repository System do you have experience? _____

Choose an option that best describes your view on the given statement.

The workflow system performs the basic tasks of managing a file entry efficiently

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree

The administrator mode offers enough functionality (editing, viewing, accepting and rejecting)

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree

Navigation between different tasks (next process, returning to previous process) was easy.

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree

I understood what to do at each step of the workflow process

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree

The steps of submission are in an intuitive order.

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree

Compared to similar systems, the workflow offers similar functionality.

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree