9th International Symposium on Electronic Theses and Dissertations
IXe Symposium international sur les thèses et mémoires électroniques
June 7 – 10 Juin, 2006, Quebec City / Québec, Canada

A workflow system for ETDs. A supportive feature of the electronic publishing chain.

Eva Müller; Frida Sandgren

Electronic Publishing Centre, Uppsala University Library
eva.muller@ub.uu.se; frida.sandgren@ub.uu.se

ABSTRACT
In this paper we describe the newly developed workflow system – DIWA (Digital Workflow Application) - which is designed to facilitate the pre-publication process of e-theses for authors as well as for faculty and administrative staff. We highlight why and how we have proceeded in the process of gathering user requirements, as well as describing the technical developments. However, emphasis is placed on examples of use and the expected effects of the DIWA on the various actors in the production and information chain.

Keywords: workflow system, pre-publication process, undergraduate theses

INTRODUCTION
How can we facilitate the management and publishing of undergraduate theses at large universities, as well as making the process smooth, efficient and rational? How can we help to improve the pre-publication process and make it more reliable? It is essential for universities of all sizes to discuss these and similar issues; it is not merely a concern for larger universities.

Uppsala University in Sweden is one of the largest Scandinavian universities with over 40 000 students annually. Since 2000, the majority of doctoral theses are available online via the DiVA publishing system[1], developed and maintained by the DiVA Project within the Electronic Publishing Centre at the Uppsala University Library[2]. DiVA offers both publishing services and technical solutions for local repositories and is used today at an additional sixteen universities across Scandinavia.

Currently more than 70% of the 400 doctoral theses produced every year at Uppsala University are published electronically through the DiVA system and a common workflow, supported by the current system, has been adopted by all doctoral students.

However, when it comes to undergraduate theses the situation is different. Every year there are several thousand undergraduate theses produced at Uppsala University. In contrast to doctoral theses, these are managed at faculty level by the supervisors, administrators and examiners of the different educational programs. Since the practices adopted by various university departments differ somewhat, there is a need for more flexibility. This is also mirrored in their requirements as concerns a workflow system.

[1] For more information about DiVA, see http://www.diva-portal.org
[2] http://en.wikipedia.org/wiki/Uppsala_University_Library

DIWA – Digital Workflow Application - is a workflow system for electronic publishing currently under development at the Electronic Publishing Centre at the Uppsala University Library. The main portion of DIWA is being developed within the DiVA project, though two other external bodies[3] have provided additional resources. In 2005, as the first step in the development of a new workflow system, a working group within the DiVA consortium was initiated with the aim of identifying system and user requirements for the development of a more flexible and enhanced workflow for publishing of undergraduate theses in DiVA. This group carried out a survey with the aim of examining the management of undergraduate theses at different departments and faculties. The work of the group resulted in a specification for a new workflow system – DIWA. DIWA is intended to be an open source application and will also serve as a new component of the next version of the DiVA system – DiVA 2.

METHODOLOGY

One of the main challenges of this project has been to establish effective and efficient methods for understanding user needs and, in the next step, for expressing these in a fashion that allows efficient communications with systems developers. The methods used to support this are those typically applied in user-centered design such as focus groups, interactive surveys and scenarios. The focus group consisting of system users from five Swedish universities has developed the survey jointly. The survey incorporates detailed interviews with faculty and administrative staff at five Swedish universities. These descriptions of current workflows and local practices for management of undergraduate theses have helped to specify desired user requirements. The results of the interviews were modeled as use cases and later translated into UML use case diagrams. In the requirement specification, both functional and user requirements are expressed as scenarios or paths through the separate use cases.

User requirement specification

Essential use cases in the specification mirror the difficulties that the undergraduate thesis management and reviewing processes often meet. For example, the reviewing process often involves several actors (e.g. student, opponent, supervisor and examiner) all of whom need access to the thesis at different stages in the process. Other use cases reflect requirements such as document version checks, e-mail and communication areas, digital storage of files etc.

DIWA – Digital Workflow Application

As proved by the results of the survey, one recurring issue that staff at large university departments have to deal with is the actual numbers of theses to be processed, and the necessity of keeping track of the stages in the procedure for each thesis (submission, reviewing, rewriting, final submission etc). In DIWA, this problem has been solved through a set of defined states which the document must pass through in order to be accepted and published. A chain of states is created when the system is initiated, and the user roles who may deal with the different states are also defined. E-mails using templates may also be issued whenever there is a change of state for any particular thesis. Figure 1 below shows an example of a list of states, arranged from top to bottom, in a defined workflow for undergraduate theses.

[3] BIBSAM - National Co-ordination and Development: http://www.kb.se/bibsam/
and Uppsala University: http://www.uu.se

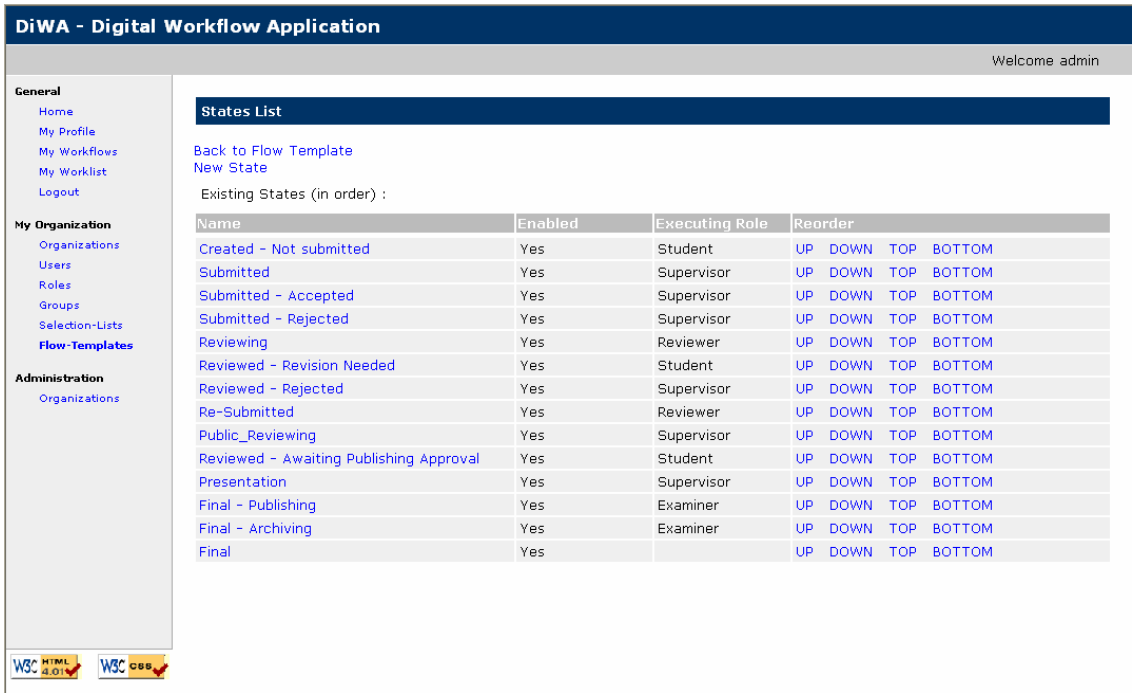

Figure 1

In DIWA, any document object (e.g. undergraduate thesis) will follow a pre-defined and customized 'path' or 'flow' based on a 'flow template'. The flow template is created when the system is initiated and constitutes the core of DIWA. It is possible to clone 'basic' default templates and add or omit workflow components such as users or user roles, organizations, states or metadata. There is also built-in support for importation of local information – e.g. selection lists of users, educational courses, subjects etc. from various sources.

Examples of use

Each user role has access to a particular view of DIWA. For instance, an administrator may have access to system and workflow settings, but authors have access to files and metadata upload areas etc. Figure 2 below shows the view for workflow settings.

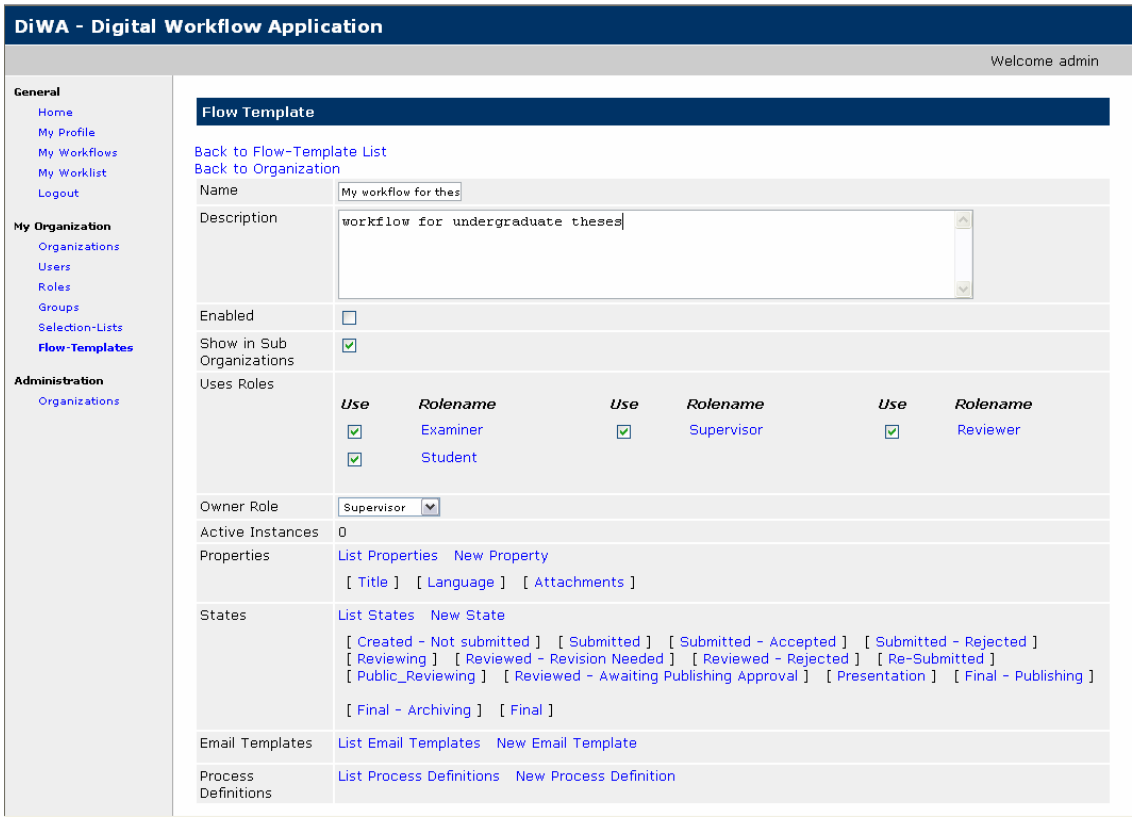

Figure 2.

Moreover, each user is able keep track of his or her assigned work through '*My Worklist*' (e.g. a list of objects that await action by the user) and refer to '*My Objects*' (e.g. user's manuscripts and their current status).

Publishing and storage

Any document object is stored along with a pre-defined set of metadata which is completed at different stages in the process. When the undergraduate thesis is completed through DIWA, it is finally published via *DiVA* and stored in an XML based format (the DiVA Document Format) in local archives and repository.

What makes DIWA different from other workflow systems?

DIWA is tailored, but not limited, to workflows for electronic publishing within academic contexts. It is suitable not only for undergraduate theses, but for any type of electronic publication which requires some process of digital reviewing, and where the interaction between users powers the workflow.

DIWA is based on open standards (Java technologies) and offers service-centric functionality. Business processes and protocols can be described through BPEL to enable task-sharing using a combination of web services. These external services can be incorporated into workflows when they are created in DIWA, building Service Oriented Architecture (SOA).

EXPECTED OUTCOME

The authors and academic staff involved in the processing of theses will reap many benefits from an advanced and flexible workflow system. The major advantage for both students and academic staff is that it will be easier and more reliable to keep track of the processing of theses. This may also exert a considerable effect on efficiency in terms of the time required, and costs, for the entire process.

Moreover, by imposing standards on the metadata and templates for document processing, the final documents will be well-structured and easy to integrate into other systems, if desired. This will bring benefits for university administrators as well as for librarians or staff running digital repositories. Librarians will be provided with facilities that enable them to easily integrate published material into library collections or local institutional repositories, as well as adding metadata information to their electronic catalogues.

Though the system has been designed focusing on the pre-production phase of the publishing process, benefits are also expected in other parts of the production and information chains.

ACKNOWLEDGEMENTS

We would like to express our gratitude to the DiVA working group for undergraduate theses, especially to Markus Schneider, Ingela Tång and Peter Berkesand, for such thorough work with the survey and for their interesting input. Also, many thanks to Swedish BIBSAM as well as to Uppsala University for additional funding. Finally, our deep gratitude to our colleague at EPC, Erik Lindqvist, for working so patiently on the design and technical development of the system.