



Vrije Universiteit Brussel

University Library
Library and Information Technology Department
Vrije Universiteit Brussel

Supporting change and diversity for theses in an institutional repository

ETD2007 - the 10th International Symposium on Electronic Theses and Dissertations

Wouter Heyse

june 2007



Abstract

In an academic institution, theses and dissertations are an important part of the learning process. Although standard approaches exist for electronically handling theses in a library, there is still a wide variety of theses to support. Some disciplines pose specific requirements on the theses (e.g. in architecture) and departments can have different policies on creating and submitting theses. The library has to cope with diversity in the thesis metadata and it should be aware that they can change over time. This paper addresses the consequences for the library and proposes a number of solutions to handle the submission and management of electronic theses. The goal is to establish a digital repository which can adapt to new developments in the thesis policies and metadata. The institutional repository at the Vrije Universiteit Brussel (Brussels, Belgium) is used as a case-study.

The deployment of digital repositories is more and more a necessary task for institutions. The internet has become an important aspect in academic teaching and research. There is a steady increase in information digitally available, new web appliances are constantly emerging and web 2.0 has created a community atmosphere around the internet which made users more flexible and demanding. The digital world is always changing, refining and redefining itself. Institutions who want to establish a repository are themselves in constant change too. New government policies in Belgium have made universities eager to measure their own teaching quality and research output, for example. Internally, multiple policies in one institution can exist (for example, in different departments). Additionally, the establishment of e-learning, course management and ERP systems can have a large impact on institutional policies. It is an increasingly challenging task to blend the repository seamlessly into these new environments.

We can divide the deployment of a repository into three parts: integration with existing infrastructure, institutional policy and data collection. The integration of the repository is important both on a usability and structural level. This means that the repository will have to handle extra requirements beyond the scope of a standard library and it will have to use databases outside the library formats. The institutional policy greatly determines the usability aspect of the repository. It defines how much control the user is given, and how visible the repository is to the community. The data collection is, for the library, the most important issue. The collected data must be consistent, correct and well-formed.

We use the institutional repository of the Vrije Universiteit Brussel (VUB) as a case-study to discuss these difficulties.

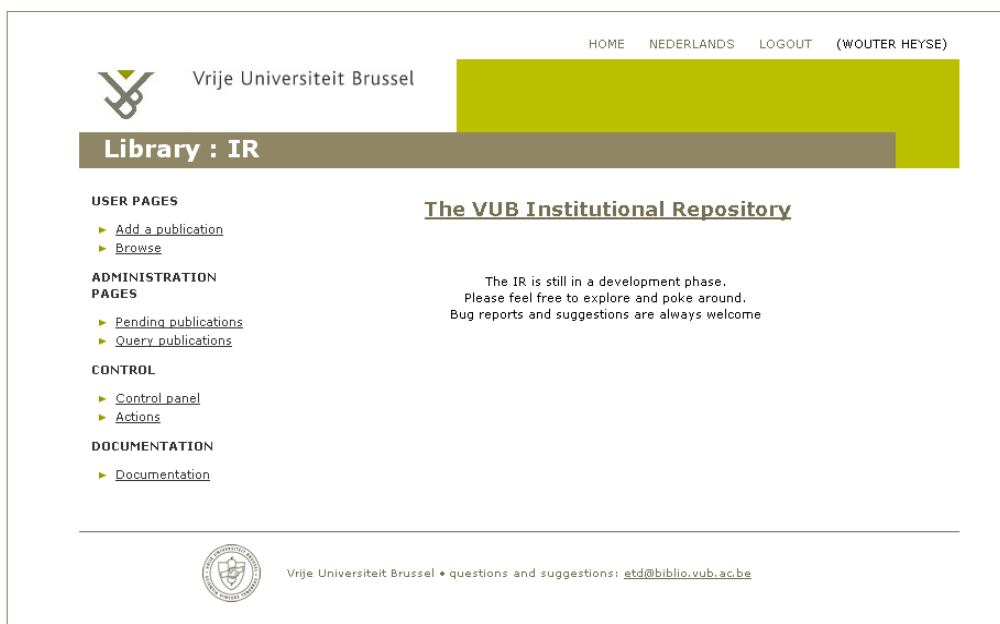


Figure 1: Screen capture: main window

The repository had to be integrated with two partners: the library catalog and the research database. This means that the repository has to adhere to the current library standards (e.g. MARC21), but it must also maintain an extensive list of extra metadata (+10 fields). This excluded the use of Dublin Core as a means for data storage.

The global academic policy at the VUB is very straightforward (phd theses are mandatory) but at the departments level, the policy is much more specific and requires an individual approach. Some departments require only the best theses to be available online, for example. Additionally, the governmental policies are changing rapidly. The government is adjusting the educational landscape to the European policies. This has its effects on organisational, educational and research levels. For the repository, it means that the very definition of a thesis can change, and that institutions are becoming more flexible in the various means of awarding degrees.

The collection of data is the most important task of the repository and requires some attention. As already mentioned, our repository has quite a lot of metadata fields. It is therefore necessary to track what metadata is

included in the document types (thesis, paper, ...) and in which format. The metadata can change over time too. For example, a department might require to store the thesis grade, or a phd thesis could have two instead of one promoter. Additionally, new document types could emerge (a bachelor thesis, for example).

It is clear that we need a flexible, maintainable system to support these requirements. Therefore, we have developed a prototype in an attempt to resolve them. The goal is to build a modular repository system, so that changing the repositories functionality is both flexible and easy. The result is a highly configurable web interface, where its look, behaviour and storage can be edited easily (see figure 2).

The prototype is designed to handle every aspect of the website dynamically. The existing web interface is extended with configuration controls. These controls enable the maintainer to add new document types. This is done by defining a data model for the metadata and several views for the interface. The content of the views can be defined and changed dynamically by inserting dynamic content objects (eg. a text field, a table, ...). These objects can be shaped by adding text and specifying configuration options. There are extra tools included to maintain and manage the interface of the repository (see figure 3)

The main advantage of the prototype is that supported features in the interface can be added and changed very easily. There is no need for an extensive use of configuration files or code snippets to make the website work. Because the configuration controls are generated within the interface, problems or bugs can be easily located and fixed. However, as with all generated code, extending the code base with additional functionality requires more attention and time than with a static html website.

■ Step 1: General

Title *: ?

✕ Abstract *: ?

✕ Publication date *: ?

✕ Language: ?

* this is a required field

✕ ✕

Figure 2: Screen capture: close up with configuration controls

abstract (textarea)

info attributes data texts type edit

■ Type
textarea

■ Path
"paper","fields","stage1","fields","titlesection","fields","abstract"

■ Navigation
Parent: [titlesection](#)
Contents:

Figure 3: Screen capture: the configuration interface

Some difficulties still remain in the prototype. Foremost is it impossible to predict (or guess) which additional requirements will emerge for the repository. In the prototype, we have made an effort to support new and changing document types and fields. However, incorporating business logic into the repository is far more complex to handle in a configurable and maintainable way.

As a conclusion we can state that supporting change and diversity in a repository is a challenge. There is a need for a repository system which can handle a constantly changing digital environment. The presented prototype enables the maintainer to adjust the repository to the current institutional requirements. By supporting data collection in a configurable and extendible way, the prototype greatly reduces the cost of maintaining the repository system.