
# Proposal Document

## *Using DSpace Open Source Software to implement a Digital Repository at the University of Pretoria*

*University of Pretoria*
*Pretoria*
*0001*
*http://www.library.up.ac.za/edu/dspace*

*Tel.: +27 12 420 3082*
*Fax: +27 12 362 5100*
*E-mail: ina.smith@up.ac.za*

# Document Information

**Compiled by:**

| *Name* | *Date* | *Signature* |
|---|---|---|
| Ina Smith | 20 October 2004 | |

**Reviewed and edited by:**

| *Name* | *Date* | *Signature* |
|---|---|---|
| Dr Heila Pienaar | 22 October 2004 | |
| Dr Heila Pienaar | 25 October 2004 | |
| Maureen du Pisanie | | |
| Rita Badenhorst | | |
| Sanet Haupt | | |
| Monica Hammes | | |
| Amelia Breytenbach | | |
| Ria Groenewald | 1 November 2004 | |
| Christel Smith | 29 October 2004 | |
| Leonard Daniels | | |
| Prof Theo Bothma | | |

**Approved by:**

| *Name* | *Date* | *Signature* |
|---|---|---|
| Dr Heila Pienaar | | |
| Prof Hans Boon | | |

# CONTENTS

# 1 INTRODUCTION

A digital institutional repository in the context of a higher education institution can be defined as *"a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution."* (*Lynch*, 2003). It was decided by the E-Information Management Team that the library in collaboration with Education Innovation will initiate this digital repository project at the University of Pretoria (UP): *"At the initial meeting of the e-information workgroup Prof Boon stressed that this will be a university project and emphasis must be on the preservation of information"* (*Metadata Workgroup*, 2004).

An evaluation of available open source software platforms were conducted in order to make an informed decision on the most appropriate platform which will fulfill in the needs of the UP community (i.e. faculties, departments, centres, staff, students, focus areas etc.). The following platforms were reviewed: Arno, CDSware, dSpace, ePrints, Fedora, MyCore, I-Tor. *Click* to view Evaluation. Due to a lack of time and other resources the platforms were only reviewed from the end-user's viewpoint.

The four specialized digital focus areas that have been identified for inclusion in the digital repository at the University of Pretoria, are:

- Africana Collection
- Digital Learning Objects
- Scholarly Communication
- Arts & Culture Digital Objects

This proposal addresses the following details for establishing a Digital Repository at the University of Pretoria, using the DSpace platform:

- Project definition
- Approach
- Project phases and deliverables
- Project plan
- Assumptions

# *2 PROJECT DEFINITION*

**Objective**

The objective of this project is to:

- Develop a digital institutional repository for the University of Pretoria, using DSpace;
- Identify services that will be made available via DSpace;
- Formulate general policies and guidelines for submitting digital objects to DSpace;
- Customise DSpace according to the needs of clients and the four identified focus areas.

**Scope**

The scope of this project is to:

- Conduct an analysis of user requirements;
- Structure Collections and Communities within DSpace in collaboration with representatives from the various focus areas;
- Identify metadata elements as applied to the various Collections in collaboration with representatives from the various focus areas;
- Structure the workflow within each Collection in collaboration with representatives from the various focus areas;
- Provide ongoing training and support to clients;
- Host and preserve materials within faculties;
- Supply system monitoring, back-up and recovery;
- Offer a web-service which will provide access and support to clients;
- Internally market the new system at UP and the library;
- Report on progress to the Digital Repository Management Committee, and to the Library Management Team;
- Identify system enhancements, and integrate future system developments by MIT and HP.

**The following is not in the scope of this project:**

- External marketing;
- Negotiating for required hardware and software to continue with this project;
- Policies, procedures and project plans for each specific Collection (will have to be compiled by the coordinators of the four focus areas).

**Target groups**

The target groups ("clients") for this project will be the UP Community – i.e.:

- Present and future staff members of the Department of Library Services, University of Pretoria;
- Past, present and future academic personnel at the University of Pretoria;
- Focus areas, faculties, centers, institutes and research divisions at UP.

**Strategy**

The strategy for the development of this project will be:

- Analysis: A needs analysis will be conducted to be sure that the needs of the UP community are addressed;
- Design: This implies a structural design, rather than a graphical design. Designing and customisation of the DSpace open source software application in alignment with the E-strategy of the library and UP;
- Development: Development of the open source software application to address the needs of the UP Community, in collaboration with the international DSpace Federation, within the structure decided upon;
- Implementation: Implementation of the open source software by uploading digital objects identified by the four focus areas;
- Evaluation: Evaluation (Summative and Formative) of the institutional repository to be sure that the needs of the UP community are met.

# *3 APPROACH*

**Analysis**

A needs analysis will be conducted to be sure that the needs of the UP community are addressed, and also to identify deficiencies and address tools and features the community can benefit from. This will consider the main objectives of the Library E-Information strategy, i.e.:

- To take part and make a contribution to the international e-information phenomenon, e.g. open access, digital preservation, eScience, content management;
- To support education innovation and research excellence at UP;
- To deliver optimal e-information portal services (workflow) to our clients (*E-Information Strategy Workgroup*, 2004).

**Designing the structure of the digital repository**

DSpace will be accessible via the library web-page at
*http://www.library.up.ac.za/*
In designing the digital repository we will make use of the submission, digitisation and metadata standards compiled by the various workgroups. We acknowledge the fact that these documents are still "work in progress".

We plan to use a sectional structure in the digital repository where the content will be presented as different Top-level Communities, Sub-Communities within the Top-level Communities, and Collections within the Sub-Communities. An example of this structure is illustrated on the following page:

**Example:**

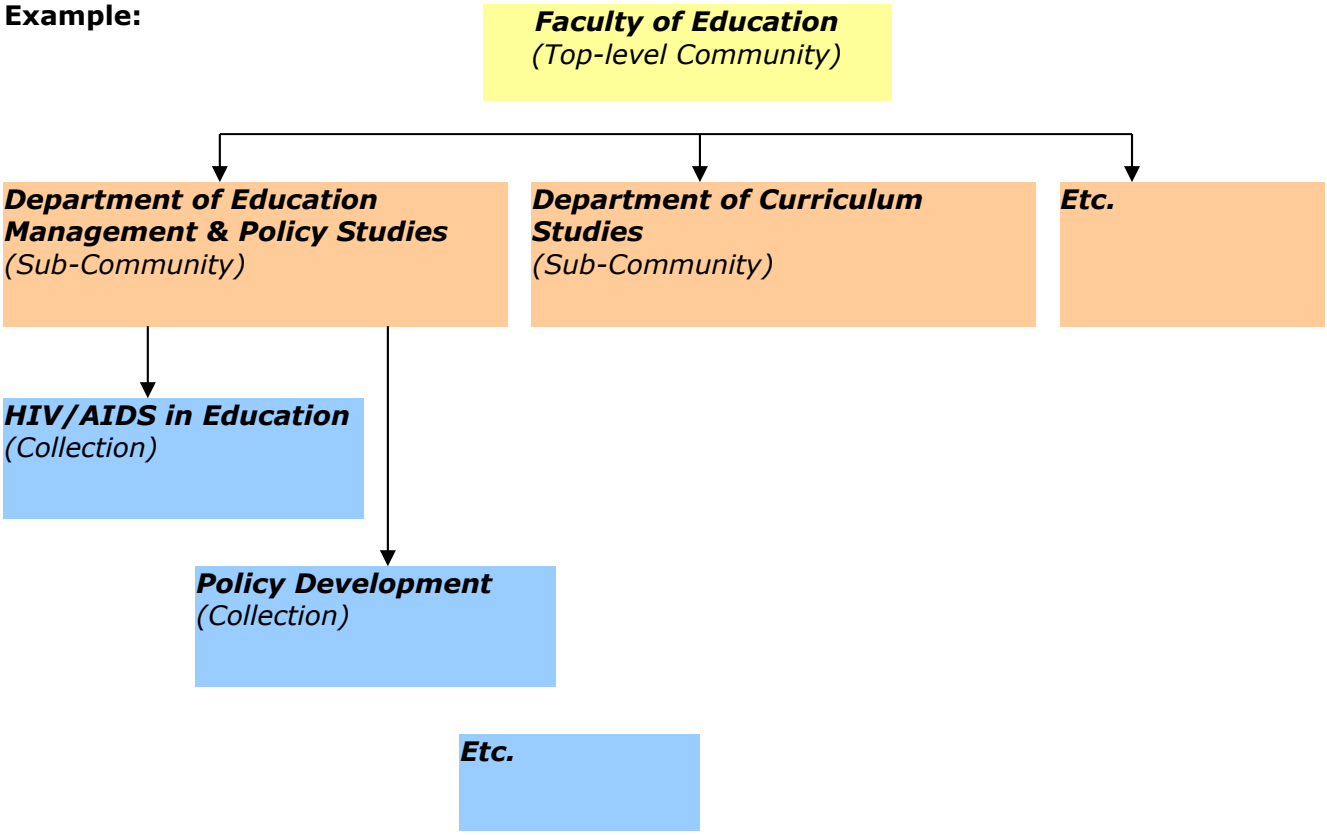


*Figure 1: Example of a possible structure within DSpace*

The following levels of involvement are foreseen by the parties involved:

- Provision of information and digital objects to be uploaded into DSpace;
- Participating in quality assurance of the digital objects and testing of the open source software.

The following digital objects were identified (*Pienaar*, 2004) to be submitted to the various Collections within DSpace:

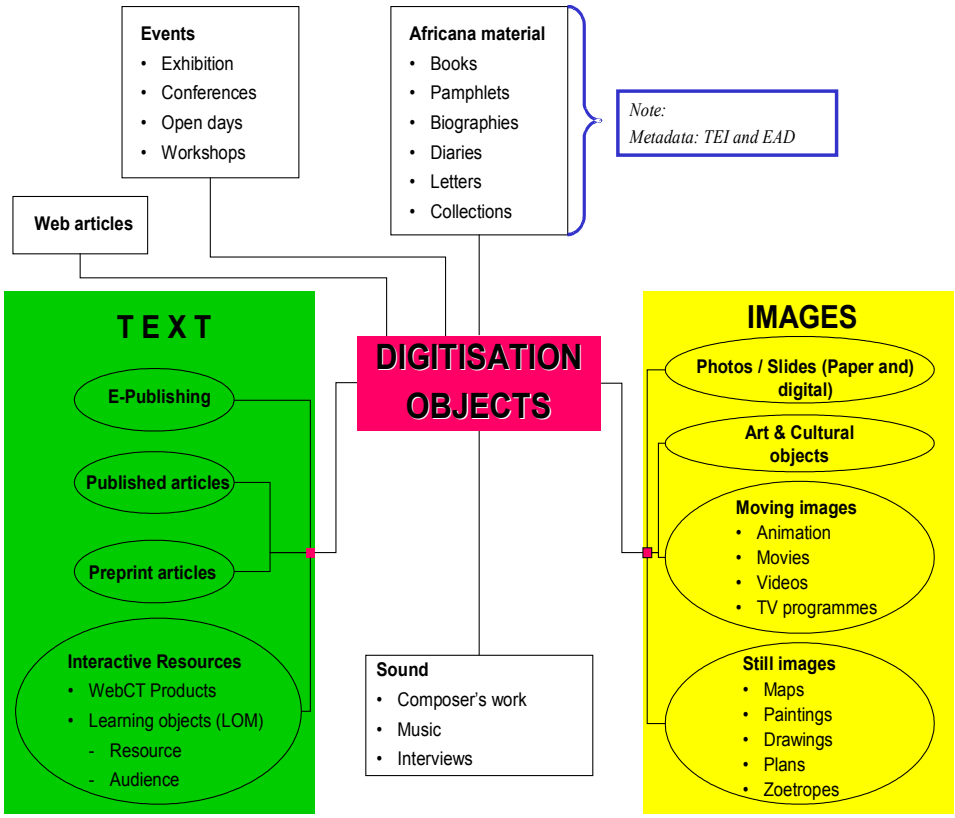


*Figure 2: Digitisation Objects*

## Development of the digital repository

The development of the digital repository will be guided by the following:

- Existing Collections, Centres, Schools, Faculties and Departments within the structure of the University of Pretoria;
- The following focus areas that were identified: Africana Collection, Digital Learning Objects, Open Access (Scholarly Communication), Arts and Culture Digital Objects;
- Digital objects submitted to the institutional repository will include the following (*View Figure 2*):
    - Text (E-Publishing, Published articles, Preprint articles, Interactive resources)
    - Events
    - Web articles
    - Africana material
    - Sound
    - Images

**Services**

Services that will be made available via DSpace (*Barton and Walker*, 2002):

- Submission services
- Access Services
- Community Management Services
- User Support Services
- Core Repository Services
- System Management Services

**Management, Staffing, IT Support and Training**

**Management Team**

The Management Team will be responsible for the following:

- Management of UP Digital Repository;
- Formulation of policies related to system enhancement and direction;
- Establishing procedures;
- Marketing of new product;
- Management of platform;
- Training of clients;
- Negotiating for funding and other resources;
- Consultation with clients (Pienaar, 2004).

**Co-ordinators**

The co-ordinators of the four focus areas identified will be responsible for:

- Compiling their own teams from existing library staff, contracting time and availability of team members;
- Compiling project plans;
- Formulation of procedures and policies regarding specific Collections and Communities;
- Recommendations regarding system enhancements and direction, requirements for communities and contributors, content, its nature and intended use in the DSpace environment, service offerings;
- Selection criteria;
- Metadata application (for retrieval);
- Workflow;
- Recommendations regarding services, standards, functionality;
- Promote and market DSpace concept and system within UP community.

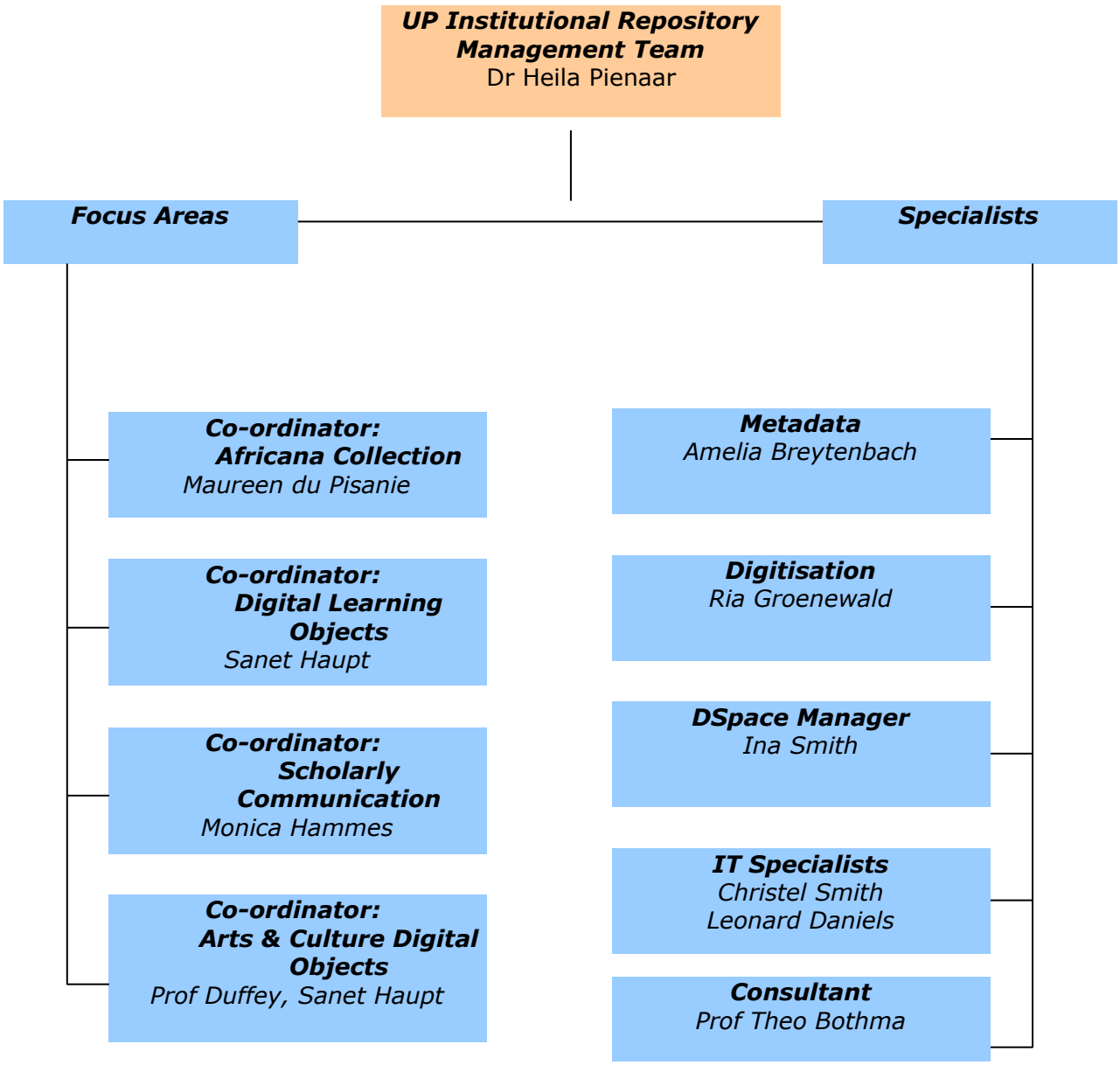


*Figure 3: UP Institutional Repository Management Team*

In order to implement and provide DSpace services with a reasonable chance of success, staff will be required. In implementing a new system and service, we anticipate the need for considerable effort to fine-tune both the technology and the service component in the first year. In addition, we expect considerable growth in the user base. We believe that trying to accomplish the set goals without dedicated staff would be very risky. If dedicated staff are not assigned, full responsibility for the technical support of DSpace will fall to the Libraries' System's Division, and the responsibility for user support will become the work of the Information Specialists throughout the library. While we do expect Information Specialists to be involved with users, we believe that relying totally on existing staff in various units would be far too fragmented for the launch of a very important new service that needs concentrated attention. In addition, it would seriously compromise other important public services initiatives by drawing away the time of staff members responsible for their planning and implementation. As DSpace expands in scope and functionality, management and support efficiencies will have to increase simultaneously.

**DSpace System Manager (Leonard Daniels)**

The DSpace System Manager should have the following skills:

- Technical management skills;
- System monitoring, testing, debugging;
- Develop portions of DSpace related to system administration;
- Monitor and upgrade DSpace utility programs and middleware;
- Develop approved system enhancements;
- Manage hardware contracts and system administration;
- Java programming;
- Networks;
- Unix/Linux Server.

**DSpace User Support Manager (Ina Smith)**

The DSpace User Support Manager will be responsible for:

- Client Support Training;
- Coordinate and manage the definition and setup of new DSpace communities;
- Plan and implement usability tests;
- Make recommendations on new functionality for DSpace;
- Chair the DSpace Workgroup;
- Write and maintain user documentation for the system, help pages;

Support will be made available via the DSpace web-page at *http://www.library.up.ac.za/edu/dspace* or e-mail *upspace@up.ac.za*

**Web Manager (Ina Smith)**

- Apply usability and user interface design knowledge and expertise;
- Integrate DSpace into UP web-sites.

**Metadata Specialist (Amelia Breytenbach)**

- Share knowledge and expertise about Qualified Dublin Core (as implemented by DSpace) and consult with the DSpace User Support

Manager on questions, issues related to the MARC to Dublin Core metadata crosswalk;
- Adaptation of metadata elements in DublinCore registry of DSpace according to Metadata Standards document;
- Provide training.

**Digitisation Specialist (Ria Groenewald)**

- Serve as knowledge expert for digital preservation issues;
- Provide training;
- Share knowledge and expertise on matters of archival selection, preservation and UP records policies;
- Provide advice regarding scanning of paper or microform documents to digital formats, as well as reformatting of documents already in a digital format.

**Information Specialists**

Provide general information about DSpace as a service of the library and UP;
Alert users to the information potential of the repository;
Assist end-users with searching the repository;
Answer end-user questions about DSpace;
Provide information about the possibility of contributing to the repository.

**Systems Division**

Serve as backup for the work of the System Manager. Provide some technical support for the DSpace platform and application, and integration with other library systems.

**Meetings**

The Management Team will meet once a month. Members of various focus areas will meet quarterly. Constant interaction between the Management Team and Focus Areas will guide the process.

**Training**

Different levels of training will be provided to different groups within the library and UP. A detailed training program will be made available at a later stage.

*All Library Staff*
Basic understanding and use of DSpace, its mission and services

*Information Specialists, Coordinators & Representatives of focus areas*
In-depth DSpace Training

*Cataloguing Staff*
DSpace Metadata, DublinCore

*Systems Division*
DSpace Systems Training

*UP Faculty Community*
Introduction and basic understanding of DSpace. Training will be conducted by Information Specialists.

# *4 PROJECT PHASES AND DELIVERABLES*

The tasks and deliverables for each phase are listed in the following tables:

| **1. Needs Analysis** | **Deliverable** |
|---|---|
| Attend kick off meetings | None |
| Needs Analysis<br>Resource Analysis<br>Conduct structural requirements assessment<br>• Approach<br>• Schedule<br>• Strategies<br>• Logistical and resource requirements | Needs Analysis Report |
| Sign-off | |

| **2. Design** | **Deliverable** |
|---|---|
| Structural design of institutional repository:<br>• Top-level Communities<br>• Sub-Communities<br>• Collections<br>Description for each Community/Collection | Design Blueprint |
| Workflow within each Collection | Workflow Document for each Collection |
| Interface design | DSpace Interface |
| General policies and procedures, and Community specific policies and procedures | General Policy Documents and specific policies for each Community within four focus areas |
| Training material | Training Document |
| Metadata standards | Metadata Standards Document |
| Digitisation standards | Digitising Standards Document |
| Submission standards | Submission Standards Document |
| Sign-off | |

| **3. Development** | **Deliverable** |
|---|---|
| DSpace Web Page with Help & FAQ | Web page |
| Installation/Configuration of software | UPSpace |
| Incorporate specifications set in documents compiled during Design | UPSpace |
| Registering e-persons & administrators | UPSpace |
| Sign-off | |

| 4. Implementation | Deliverable |
|---|---|
| Conduct Marketing | E-newsletter, Faculty mailing lists, brochure, web pages, etc. |
| Library Orientation (demo) | Presentation |
| Faculty Orientation | Presentation per Faculty |
| Training | Training Schedule |
| Sign-off | |

| 5. Evaluation | Deliverable |
|---|---|
| Review | Summative evaluation report |
| Summarise and review lessons learned from the project<br>Planning for long-term server, storage and network requirements | |
| Sign-off | |

***Launch Event***

# *5 ASSUMPTIONS*

- A DSpace System Manager will be made available for the duration of the project. This person must have a good understanding of open source software, Java programming language, networks etc.
- A DSpace User Support Manager will be made available for the duration of the project.
- Documentation compiled by the four focus areas will be made available to the DSpace Management Team (policies, workflow procedures, project plans, selection criteria).
- The following will need to be addressed in future:
    - Marketing Strategy and Plan;
    - Server availability;
    - Security;
    - Greenstone to DSpace (Batch importing – Premium Services offered by DSpace);
    - Theses and dissertations from Virginia Tech software to DSpace;
    - Metadata adaptations;
    - Submission standards.
- Policy issues regarding DSpace will need to be addressed by the Management Team e.g.:
    - DSpace Collections (responsibilities, rights, definition of a community);
    - Withdrawal of items from DSpace;
    - Provenance;
    - File format policy;
    - Authentication policy;
    - Preservation policy;
    - Minimum Metadata Requirement (See *Metadata Standards*);

- Licensing & Copyright, Intellectual Property Right Issues.

# 6 PROJECT PLAN

The project plan will be aligned with the global project plan for the E-learning strategy at the library and University of Pretoria:

**2005**

| **Key Activity** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Planning* | | | | | | | | | | | | |
| *Analysis* | | | | | | | | | | | | |
| *Design* | | | | | | | | | | | | |
| *Development* | | | | | | | | | | | | |
| *Implementation* | | | | | | | | | | | | |
| *Evaluation* | | | | | | | | | | | | |
| *Support* | | | | | | | | | | | | |

# 7 RESOURCE COSTS

The following cost breakdown is based on the information currently available.

Cost is based on the design, development and delivery of the institutional repository.

The cost per resource type is indicated in the following table:

| **Resource Type** | **Man hours** | **Rate (R/hour)** | **Cost** |
|---|---|---|---|
| DSpace System Manager | Semi-Full time | Level 9 | Max R 153 846 p.a.<br>R 199 270.39 Cost per Company |
| DSpace User Support Manager & DSpace Project Manager & Web Manager | Semi-Full time | Level 9 | Max R 153 846 p.a.<br>R 199 270.39 Cost per Company |
| Metadata Specialist | Current Staff | N/A | N/A |
| Digitisation Specialist | Current Staff | N/A | N/A |
| 4 Cataloguers<br>(1 per Focus Area) | Current staff | N/A | N/A |
| 4 Information Specialists<br>(1 per Focus Area) | Current staff | N/A | N/A |
| Java Programmer | 40 hours | R 400,00 p/h | R 16 000,00 |
| **TOTAL** | | | **R 414 540,78** |

# 8 INSTITUTIONAL REPOSITORY OUTLINE

To be compiled and submitted later.

# 9 PROJECT STANDARDS

### Metadata Standards

DublinCore open source metadata software has been identified as the metadata standard for the digital repository, and will be customized according to the needs of the various focus areas (*Metadata Workgroup*, 2004). The DSpace platform metadata fields will also have to be adapted according to the requirements set in the Metadata Standards document. *Click* to view the Metadata Standards document as compiled by the Metadata Workgroup (Metadata Specialist: *Amelia Breytenbach*).
The necessary changes will be introduced to *UPSpace* once the necessary resources become available (e.g. Java Skills).


*Figure 4: Relation between DSpace Metadata Elements, UP Metadata Standards, DublinCore Metadata Element Set*

### Digitisation Standards

Digitisation Standards have been identified by the Digitisation Workgroup and will have to be adapted according to the needs of the various focus areas within DSpace. *Click* to view the digitisation standards compiled by the Digitisation Workgroup (Specialist: *Ria Groenewald*).

### Submission Standards

General standards for submitting digital objects to UPSpace will be established by the *UPSpace* Workgroup and Management Team. It will be the responsibility of each focus area to establish their own policies, procedures, and submission standards with support from the *UPSpace* Workgroup. *Click* to view *Setting Up Dspace Communities* compiled by MIT Libraries.

# *10 REFERENCES*


Barton, Mary R. and Walker, Julie Harford. (2002). *MIT Libraries' DSpace Business Plan Project: Final Report to the Andrew W. Mellon Foundation.* Retrieved October 20, 2004, from *http://dspace.org/implement/mellon.pdf*

Digitisation Workgroup. (2004). *Digitising standards*. Retrieved October 20, 2004, from *http://www.ais.up.ac.za/edu/dspace/scanning.htm*

E-Information Strategy Workgroup. (2004). *Besigheidsplan vir die Akademiese Inligtingsdiens (AI) se e-inligtingstrategie: Junie 2004-Desember 2004.* Retrieved October 20, 2004, from *http://www.ais.up.ac.za/edu/dspace/Besigheidsplan.doc*

Lynch, Clifford A. (2003, February). Institutional repositories: essential infrastructure for scholarship in the digital age*. ARL, 226.* Retrieved October 20, 2004, from *http://www.arl.org/newsltr/226/ir.html*

Metadata Workgroup. (2004). *Open source software*. Retrieved October 20, 2004, from *http://www.ais.up.ac.za/edu/dspace/metadata.doc*

Pienaar, Heila. (2004). *Raamwerk vir 'n 'UP digital repository'.* Retrieved October 20, 2004, from *http://www.ais.up.ac.za/edu/dspace/Projekvoorsteldigrep.doc*

*Setting up DSpace Communities.* (2003). Retrieved October 20, 2004, from *http://dspace.org/implement/setting-up.html*



*I would like to acknowledge the help, contributions and hard work done by the following individuals:*
*Dr Heila Pienaar, Amelia Breytenbach, Ria Groenewald, Charles Ceronio, Leonard Daniels, Christel Smith, Sanet Haupt, Prof Theo Bothma, Monica Hammes*

*Thank you*
*Ina Smith*