# DSPACE – INSTITUTIONAL REPOSITORY SOFTWARE

Presentation at DISA Workshop
23 July 2009
Sean Carte, Kusturie Moodley & Nwabisa Rasana

DURBAN
UNIVERSITY of
TECHNOLOGY

# INTRODUCTION

- Higher education institutions are implementing institutional repositories as the mechanism for preserving and disseminating….

- There is a number of open source software products that are available for setting up institutional repositories.

# WHAT ARE INSTITUTIONAL REPOSITORIES?

- Digital collection that captures and preserves the intellectual output of an institution.
- Authors (self-)archive their research publications.
- Content can be anything from … Academic research articles, theses and dissertations; conference papers.
- Institutionally defined.
- Scholarly purpose.
- Cumulative and perpetual.
- Open and interoperable
  - Standards and protocols
  - Compliance

# BENEFITS OF INSTITUTIONAL REPOSITORIES

- Quicker dissemination of research.

- Brings recognition to institutions researchers.

- Raises the profile of the institution…..

- Increases the international recognition of national research output.

- Increases worldwide knowledge.

**Make Your Research Global**

Open Access

**DURBAN UNIVERSITY of TECHNOLOGY**

# WHAT IS DSPACE?

- DSpace is an open source software .........build digital repositories.

- Developed by MIT libraries and Hewlett Packard.

- Preserves and enables access to all types of digital content.

# ADVANTAGES OF DSPACE

# DSpace User Interface

- User interface is Web-based.

- User-friendly and easy to use.

- Capability to search and retrieve items by using the basic or advance search options.

- Browse all items by title, author, subject or issue date.

- User interface can be used out-of-the-box or can be fully customized.

DSpace Home >

# DUT IR

Welcome to the DUT Institutional Repository

Publishers who allow the depositing of published version/PDF in Institutional Repositories :

http://www.sherpa.ac.uk/romeo/PDFandIR.html

# Search DSpace

Enter some text in the box below to search DSpace.

Go

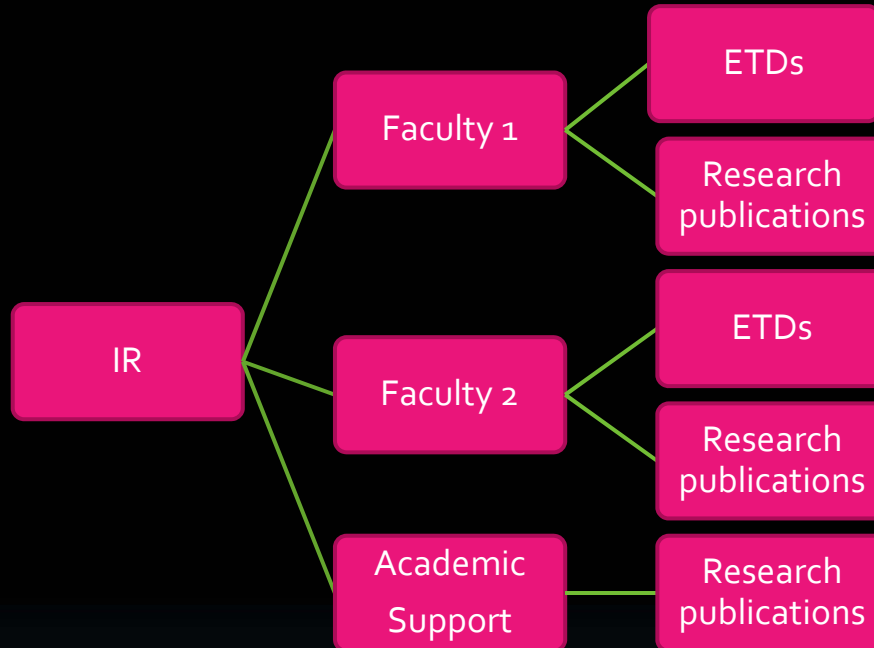# Communities in DSpace

Select a community to browse its collections.

- Academic Support
- Faculty of Accounting and Informatics
- Faculty of Applied Sciences
- Faculty of Arts and Design
- Faculty of Engineering and Built Environment
- Faculty of Health Sciences
- Faculty of Management Sciences

# DSpace User Interface

- **Content is organized into communities and collections.**

```
                                    ┌──────────────┐
                          ┌─────────┤     ETDs     │
              ┌───────────┤ Faculty 1│             │
              │           └─────────┤ Research     │
              │                     │ publications │
              │                     └──────────────┘
        ┌─────┤                     ┌──────────────┐
        │ IR  │           ┌─────────┤     ETDs     │
        └─────┤ Faculty 2 ┤         │              │
              │           └─────────┤ Research     │
              │                     │ publications │
              │ Academic            └──────────────┘
              └ Support ────────────┤ Research     │
                                    │ publications │
                                    └──────────────┘
```

- **Users to subscribe to email alerts.**

- **Full source-code is available and this enables anyone to develop add-ons.**

- **Less drastic changes are easily made.**

# Communities in DSpace

Go

**Browse**

- **All of DSpace**
  - Communities & Collections
  - By Issue Date
  - Authors
  - Titles
  - Subjects

**My Account**

- Login
- Register

Select a community to browse its collections.

- **Academic Support**

  - Research Publications (Academic Support)

- **Faculty of Accounting and Informatics**

  - Research Publications (Accounting and Informatics)
  - Theses and dissertations (Accounting and Informatics)

- **Faculty of Applied Sciences**

  - Research Publications (Applied Sciences)
  - Theses and dissertations (Applied Sciences)

- **Faculty of Arts and Design**

  - Research Publications (Arts and Design)
  - Theses and dissertations (Arts and Design)

- **Faculty of Engineering and Built Environment**

  - Research Publications (Engineering and Built Environment)
  - Theses and dissertations (Engineering and Built Environment)

- **Faculty of Health Sciences**

  - Research Publications (Health Sciences)
  - Theses and dissertations (Health Sciences)

- **Faculty of Management Sciences**

  - Research Publications (Management Sciences)
  - Theses and dissertations (Management Sciences)

# Item submission

Initial Questions → Describe → Describe → Upload → Review → License → Complete

## Describe Item

**Authors:**

[ ] [ ] Add

Last name, *e.g. Smith*   First name(s) + "Jr", *e.g. Donald Jr*
Enter the names of the authors of this item below.

**Advisor:**

[ ] [ ] Add

Last name, *e.g. Smith*   First name(s) + "Jr", *e.g. Donald Jr*

**Title:**

[ ]

Enter the main title of the item.

**Size:**

[ ]

Enter the number of pages.

**Series/Report No.:**

[ ] [ ] Add

Series Name   Report or paper No.
Enter the series and number assigned to this item by your community.

**Identifiers:**

[ISSN ▼] [ ] Add

If the item has any identification numbers or codes associated with it, please enter the types and the actual numbers or codes below.

**Type:**

Animation
Article
Book
Book chapter
Dataset
Learning Object

Select the type(s) of content of the item. To select more than one value in the list, you may have to hold down the "CTRL"

# Metadata

- Supports the Open Archive Initiative Protocol for Metadata harvesting (OAI-PMH) and other OAI-compliant archives.

- Uses a qualified Dublin Core metadata standard to identify items.

- Has a comprehensive metadata process that stores metadata in a flexible manner.

- Dublin Core registry can be extended.

- Possible to incorporate other hierarchical metadata schemes such as MAchine-Readable Cataloging (MARC) and Library of Congress Subject Headings (LCSH).

- Metadata is displayed in the item record in Dspace and is indexed for browsing and searching the repository .

# Workflows

- **Accommodate different submission workflows.**

- **Offers a sophisticated way of using the authorization policy .**

- **Different levels of authorizations can be assigned to various groups.**

Academic staff "Submitter" → Library Staff "Metadata Cataloguer" → IR Administrator "Coordinator"

# Workflows.....

- Incomplete submissions are not archived.
- Users can review their submissions for quality checking before submission.
- At any point in the submission process the user is able to stop and save the record to work on for a later date.
- If the user accidentally exits the submission process, the work is still available as DSpace automatically saves the data.
- Option to remove unwanted submissions.
- Each user who has a role in the workflow of a collection is notified of the new submission and goes to the personal workspace in DSpace to perform the assigned task.
- No limit to the length of the abstract.
- More than one file can be added.
- Administrators can administer user accounts and user groupings, create and configure communities and collections, manage files, create and modify system policies .

# Item submission

Initial Questions → Describe → Describe → Upload → Embargo → Review → License →
Complete

## Review Submission

### Initial Questions

| Multiple titles: | No |
| --- | --- |
| Published: | No |

Correct one of these

### Describe Item

| Authors: | Malunga, Philani |
| --- | --- |
| Title: | Longest solar eclipse of the 21st century |
| Type: | Article |
| Language: | English |

Correct one of these

### Describe Item

| Subject Keywords: | longest solar eclipse |
| --- | --- |

## Describe Item

**Subject Keywords:**
[                    ] [Add]
Enter appropriate subject keywords or phrases below.

**Abstract:**
[                    ]
Enter the abstract of the item below.

**Sponsors:**
[                    ]
Enter the names of any sponsors and/or funding codes in the box below.

**Description:**
[                    ]
Enter any other description or comments in this box.

[< Previous] [Save & Exit] [Next >]

---

Done

anyone yet.

## Tasks you own

| Task | Item | Collection | Submitter |
|---|---|---|---|
| *No tasks are assigned to you* | | | |

## Tasks in the pool

| | Task | Item | Collection | Submitter |
|---|---|---|---|---|
| ☐ | Awaiting reviewer's attention | An in vitro investigation of the flexural strength … | Theses and dissertations (Health Sciences) | email: Nwabisa Rasana |
| ☐ | Awaiting reviewer's attention | Perceptions on the addition of clinical practice t … | Theses and dissertations (Health Sciences) | email: Nwabisa Rasana |
| ☐ | Awaiting reviewer's attention | The influence of the workplace environment on brea … | Theses and dissertations (Health Sciences) | email: Nwabisa Rasana |
| ☐ | Awaiting reviewer's attention | Sardine run : analysis of socio-economic impact an … | Theses and dissertations (Management Sciences) | email: Nwabisa Rasana |

Take selected tasks

## Unfinished submissions

These are incomplete item submissions. You may also **start another submission** .

| | Title | Collection | Submitter |
|---|---|---|---|
| ☐ | *Untitled* | Theses and dissertations (Accounting and Informatics) | email: Kusturie Moodley |
| ☐ | *Untitled* | Research Publications (Accounting and Informatics) | email: Kusturie Moodley |

Remove selected submissions

Done

## Cost

- Zero licence fees.
- No barrier to entry - anyone can download and install DSpace on their PC....
- No outlay beyond the cost of the hardware and no recurring licence fees.

## Modest Hardware Requirements

- Hi-end hardware is not needed.
- Multiple instances can be run on the same hardware.
- Scalable.

## OS Agnostic

- Linux or MS Windows can be used for the operating system.
- One can use whatever OS one is comfortable with, provided it will run Java...

## Open Standards-Based

- Standards used by DSpace are: Java, XML, HTML (and CSS), SQL,OAI.
- Adherence to these standards ensures that data remains free, migration to competing platforms and products is possible and interoperability with other open standards-based systems is guaranteed.
- The core standard underpinning DSpace is Java, is a well-established platform.
- DSpace supports the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0 as a data provider.

## Support can be purchased

- One company (@Mire) provides paid-for support.

## Preservation

- Designed to support the long-term preservation of the digital material.

## Upgrades

- DSpace community provides frequent upgrades.

# Handles

- Handle System is administered by the Corporation for National Research Initiatives (CNRI), to assign and resolve persistent identifiers.
- This identifier does not change even if the system migrates to new hardware, ensuring constant access to the item.

# Documentation

- Extensive manual accompanies the software.
- Step-by-step instruction is provided on tasks such as installing or upgrading.

# Community Support

- A mailing list and wiki provide additional support and instruction.
- Developers of the software are active on the mailing list.
- Documentation in the wiki is provided and kept up to date by the community.

# Community Enhancements

- **Features that are not part of the core distribution can be acquired from at least one company (@mire).**
- **Usually made freely avaliable as in the case of the Elliot Metsger's Embargo add on from Johns Hopkins University.**

# Support for All File Formats

- **Can recognise and manage different file formats e.g. PDF, JPEG, Microsoft Word, etc.**

# DISADVANTAGES OF DSPACE

# DSpace User Interface

- Many customizations are not possible without the appropriate technical skills.

# Workflow

- Does not allow the user to edit a single word or letter on a field in the submission process once the field has been accepted.
- User has to delete the entire entry e.g. keyword, title, author in order to edit the field.
- The authorization system is powerful but difficult and cumbersome to use in current form.

# Open Standards-Dependent

- Along with the benefits brought by open standards there are their limitations - Web standards, particularly with HTML 4, limit what is possible in the interface.
- GUI of a thick client offers more than the CSS and form elements available to DSpace's Web client.
- Java is not everybody's cup of tea.
- DSpace does not exist in isolation - depends on a raft of other Java projects, e.g., Cocoon.

# Upgrades

- Changes that are implemented need to be documented and maintained across upgrades.
- More extensive the customisation, the harder the work involved in supporting them.

# Training

- A number of training sessions are held in the UK and USA but are limited in South Africa.
- Training must be taken care of in-house.

# Documentation

- Documentation must be understood for it to be of any use.
- Documentation appears to be written by and for people who already know the system, and probably took part in its development.

# Community Support

- Reliance on the good-will of international neighbours can be frustrating and lead to lengthy delays in sorting out problems.
- Mailing lists and Web-based resources are a fantastic and inexpensive way of trouble-shooting problems but this is time-consuming.
- Paid support is available but this is expensive.

# Community Enhancements

- Adopting an add-on generously provided by another institution does not entitle one to support.
- DSpace is constantly being developed with new versions being released periodically if one wants to take advantage of the improvements offered by new releases, one better hope that the developers of the add-ons that have been incorporated do too.

# REASONS FOR CHOOSING DSPACE

- **DUT briefly researched two potential open source repository alternatives that were in use by other South African institutions at the time: EPrints and DSpace.**

- **The University of Pretoria's effective promotion of their institutional repository, UPSpace, was a compelling reason to follow their example and use DSpace.**

- **None of the disadvantages of DSpace have been sufficiently problematic to spur any serious investigations into alternatives.**

# CONCLUSION

Many of the advantages of using DSpace recur as disadvantages. This is probably due to the fact that its greatest strength: that it is a large, complicated system can also be seen as a weakness.

Ultimately, what is really important is not the software underlying the institutional repository, but that an institutional repository exists and is utilised effectively.

# Thank you!

# Any questions?
http://ir.dut.ac.za

sean.carte@gmail.com

kmoodley@dut.ac.za

DURBAN
UNIVERSITY *of*
TECHNOLOGY