



# **Capturing knowledge in institutional repositories ... playing leapfrog with giraffes**



**Heila Pienaar (UP)**

**&**

**Martie van Deventer (CSIR)**



KM Workshop: Durban 17 August 2007





## Notes for clarification



- The giraffe must be one of Africa's best known creatures.
- We are 100% sure that giraffes do not play leapfrog!
- They do, however, symbolise those giants upon whose shoulders we know we stand!
- Leapfrogging: ...well you cannot keep on standing ... action is required ... get moving ... go play!





## Outline



- Definition & context
- Leapfrogging at the University of Pretoria
- UP Repository: capturing the University's knowledge base
- CSIR: Learning from UP
- Experimenting ... is it worth the effort
- Giraffe leap: our newly launched repository
- Knowledge CSIR 'captures'
- Tips, tricks & traps





## Definition



- In simplest terms, a digital repository is [a *database/catalogue*] where digital content, assets, are stored and can be searched and retrieved for later use.
- A repository supports mechanisms to import, export, identify, store and retrieve digital assets.
- Putting digital content into a repository enables staff and institutions to then manage and preserve it, and therefore derive maximum value from it.
- Digital repositories may include research outputs and journal articles, theses, e-learning objects and teaching materials or research data.





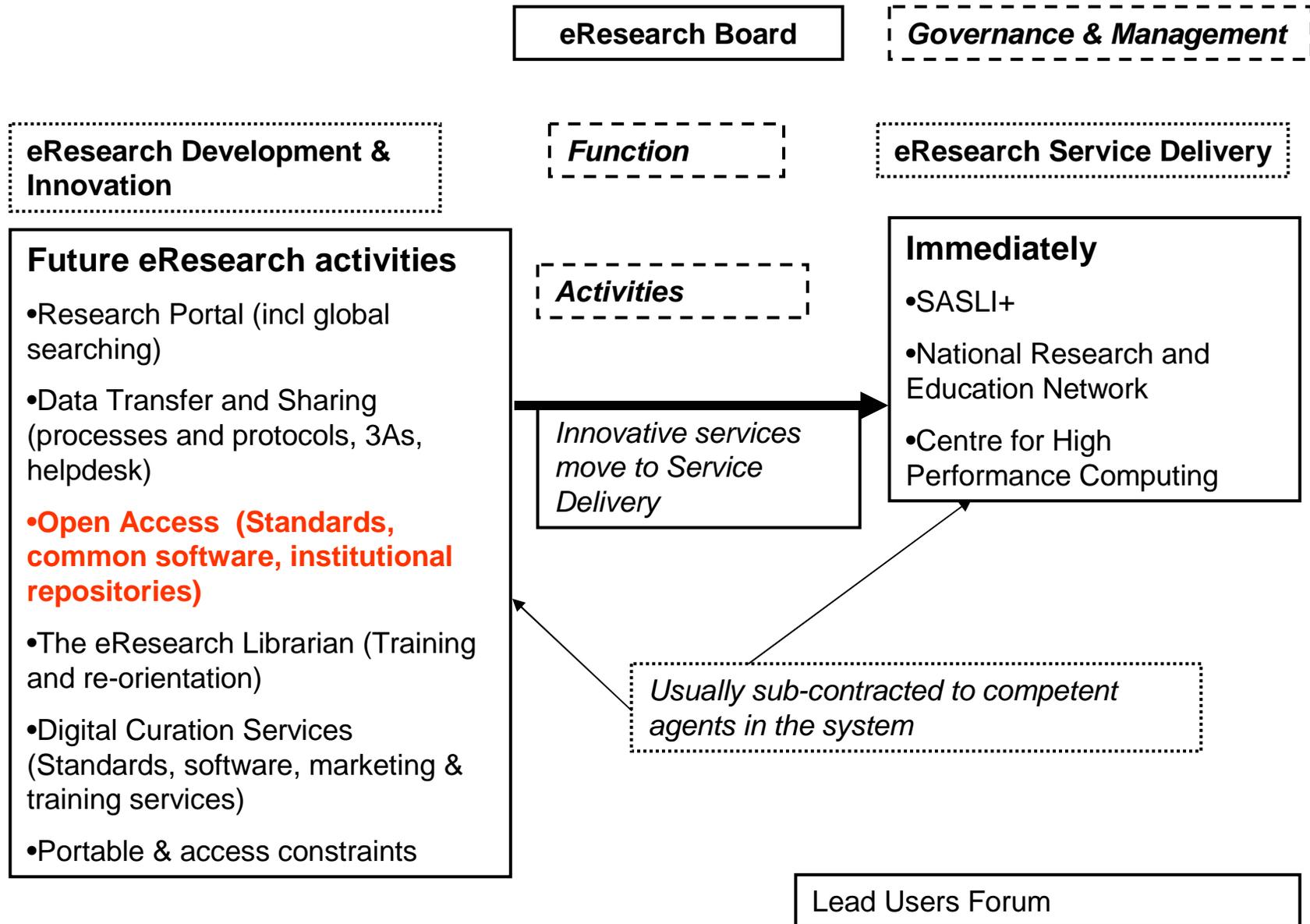
## Definition cont.



- The SPARC organisation defines institutional repositories as follows:
  - **Institutionally defined**
  - **Scholarly**
  - **Cumulative and perpetual**
  - **Open and interoperable**
- Clifford Lynch: It is most essentially an **organisational commitment** to the stewardship of digital materials, including long-term preservation where appropriate, as well as organisation and access or distribution.



# Impetus of the SARIS project: eResearch Support Service for SA





## Leapfrogging at the University of Pretoria because of existing **knowledge** on the Web & elsewhere

- 12 July 2004: Director gives assignment to dr Heila Pienaar to establish a digital repository for the University
- 23 July 2004: University wide project team assembled
- Workshops on: 2 August, 12 August, 30 August, 20 September, 4 November
- **Demonstrations of open source repository software: Greenstone, (Innovative), Fedora, e-Prints, DSpace, I-Tor etc**
- **Study OSI's "Guide to institutional repository software".**
- Choose DSpace because of UP IT fit and decentralised architecture and workflow
- Create management team for implementation of DSpace on 30 September (managers and specialists)
- DSpace web page and project plan developed: use champions
- **Metadata standards document created**
- **Digitisation standards document created**
- **Digital objects identified**
- 2005: Software development to fit DSpace into University's portal (authentication)
- Jan 2006: UPSpace launch



## WWW and other sources that made this quick implementation process possible

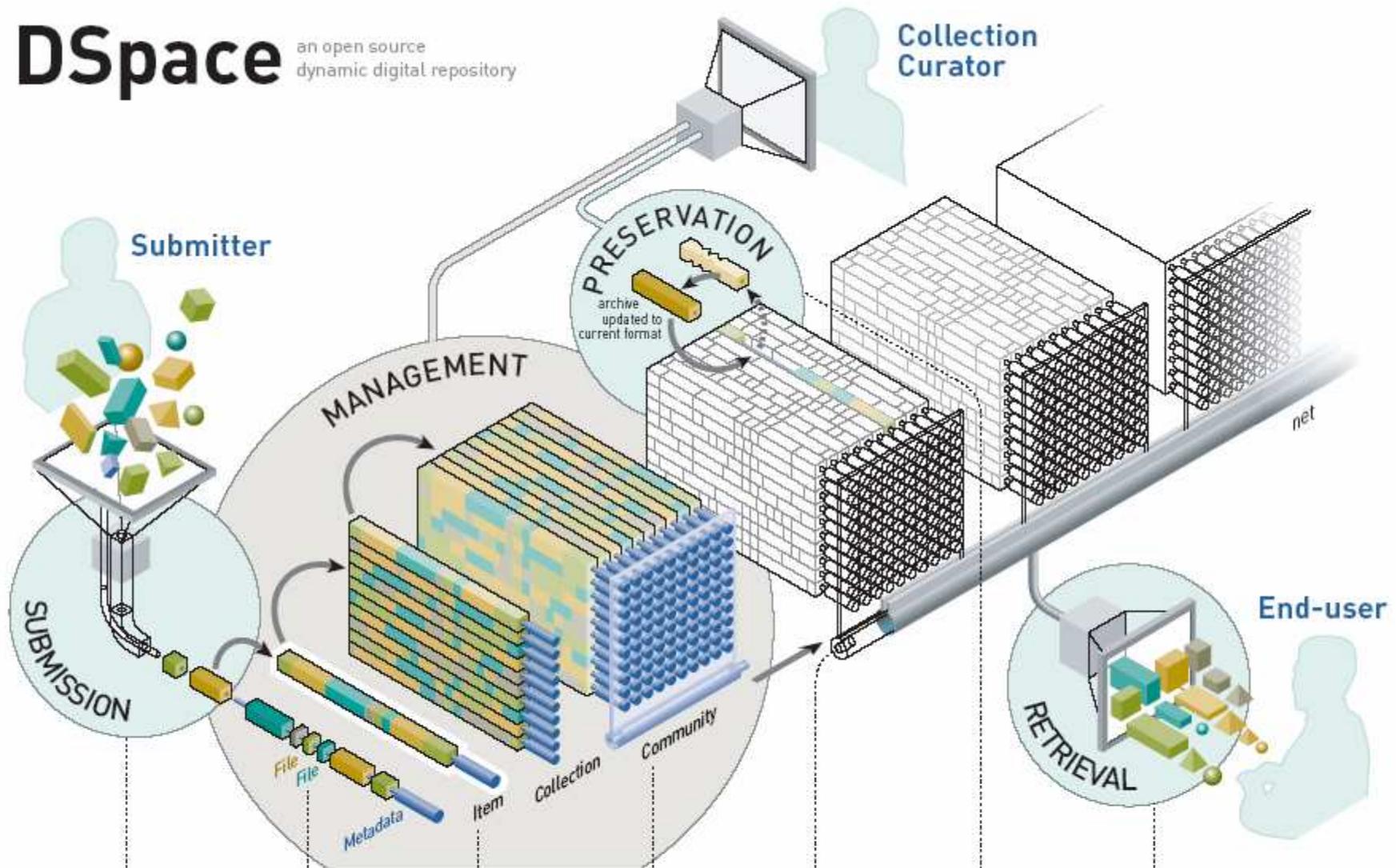


- **Open source** software available on the Internet
- Examples of the different software installations at different institutions on the Web
- Articles, presentations and discussions on the Web (**open access**)
- International courses e.g. DSpace course in India & digitisation course in the UK



# DSpace

an open source dynamic digital repository



**1** Web-based interface makes it easy for a submitter to create an archival item by depositing set of related files. DSpace was designed to handle any format from simple text documents to datasets and digital video.

**2** Data files, also called bitstreams, are organized together into bundles. Each bitstream is linked to bitstream format and encoding information. Description of each item is preserved in metadata.

**3** An item is an "archival atom" consisting of grouped related content and associated metadata. An item's exposed metadata is indexes for browsing and searching. Items are organized into collections of logically-related material.

**4** A community is the highest level of the DSpace content hierarchy. They correspond to organizational bodies in the institution such as departments, labs, research centers or schools.

**5** DSpace's modular architecture allows for creation of large multidisciplinary repositories that ultimately could be expanded across institutional boundaries.

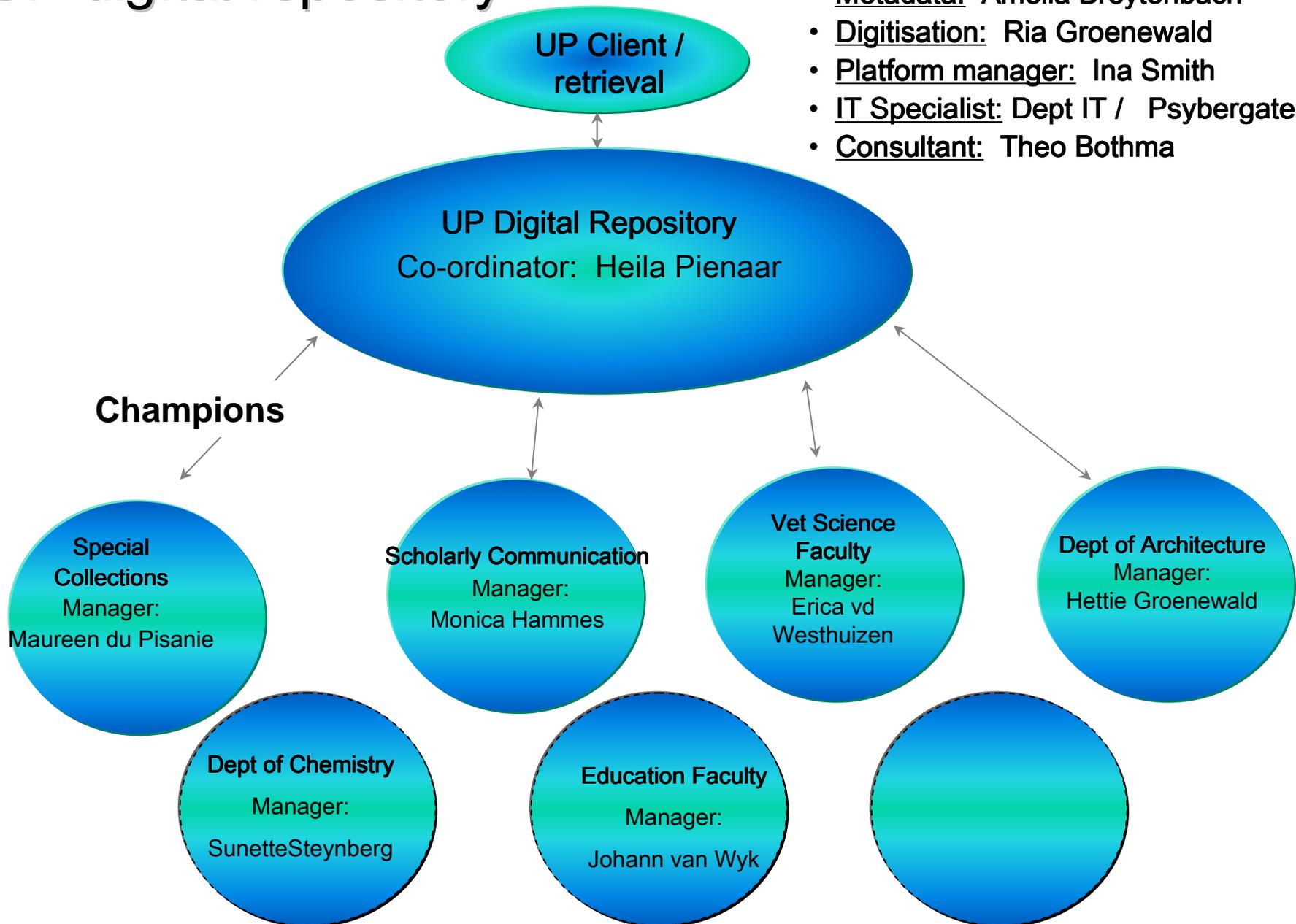
**6** DSpace is committed to go beyond reliable bit preservation and provide functional preservation where digital file is kept usable as technology formats, media, and paradigms evolve.

**7** End-user interface supports browsing and searching the archives. Once item is located Web-native formats files can be displayed directly in a browser while others can be downloaded and opened using application able to interpret the file.

# UP digital repository

## Specialist / expert roles:

- Metadata: Amelia Breytenbach
- Digitisation: Ria Groenewald
- Platform manager: Ina Smith
- IT Specialist: Dept IT / Psybergate
- Consultant: Theo Bothma



# UPSpace today! Capturing the University's knowledge base

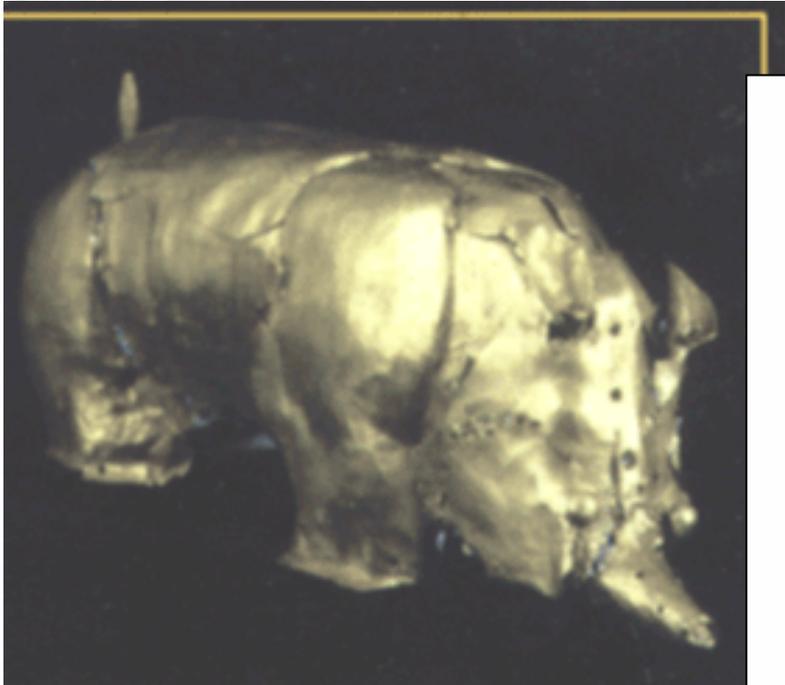
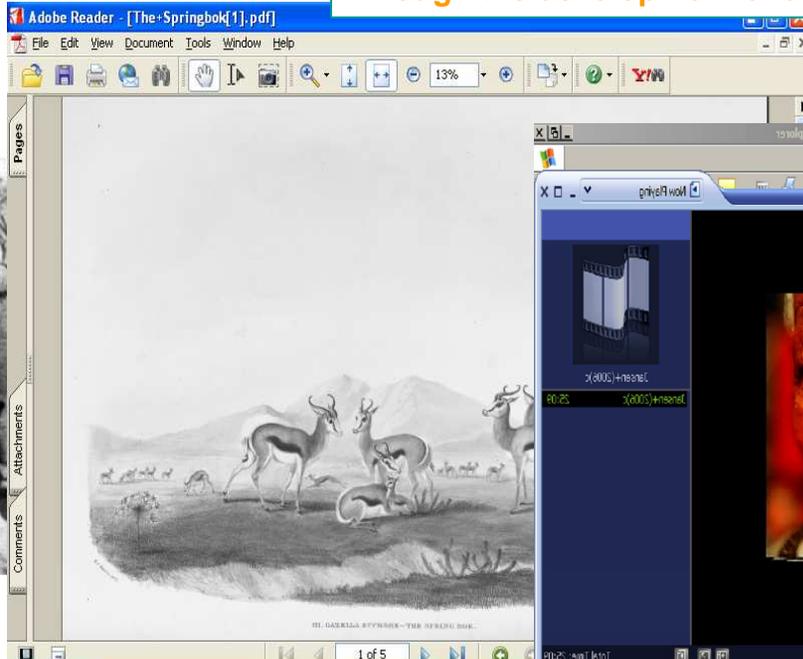
The screenshot shows the UPSpace website in a Microsoft Internet Explorer browser window. The browser's address bar displays the URL <https://www.up.ac.za/dspace/>. The website header features the University of Pretoria logo on the left and the 'UPSpace' logo in the center. Below the header, the page is organized into several sections:

- Search UPSpace:** Includes a search input field with a 'Go' button and a link to 'Advanced Search'.
- Navigation:** A 'Home' link is provided.
- Browse:** A list of navigation links: 'Communities & Collections', 'Titles', 'Authors', and 'By Date'.
- Sign on to:** A list of user-related links: 'Receive email updates', 'My UPSpace authorized users', 'Edit Profile', 'Help', and 'About DSpace'.
- Welcome Message:** A central box with the text: 'Welcome to UPSpace, the digital research repository of the University of Pretoria!' and 'Items archived since January 2006: 1 666'.
- Search Section:** A section titled 'Search' with the instruction 'Enter some text in the box below to search UPSpace.' and a search input field with a 'Go' button.
- Communities in UPSpace:** A section titled 'Communities in UPSpace' with the instruction 'Choose a community to browse its collections.' and a list of community links: 'Academic Information Service (Library)', 'Economic & Management Sciences', 'Education', 'Education Innovation', 'Engineering, Built Environment & Information Technology', 'Health Sciences', 'Humanities', and 'Law'.
- Right Sidebar:** Contains promotional text: 'Do you want to become part of the UPSpace research community?' with links to 'Contact us' and 'apply for a new collection'. It also includes 'How to submit an item to UPSpace' with links to 'View or print the pamphlet' and 'follow the animated tutorial'. At the bottom, it says 'Download Adobe Reader to access .pdf files' and 'UPSspace Web Page' with a link to 'Sign our electronic guest book'.

The browser's status bar at the bottom shows the page is from 'Home of the... ie • Innov...'. The browser window title is 'UPSpace at the University of Pretoria: Home - Microsoft Internet Explorer'.

# Examples of UP knowledge captured

For more info: Tuesday 16:00-18:00 - "Adapt or die": Energizing library staff and academics through the development of digital repositories



## The influence of high energy proton bombardment on the electrical and defect properties of single ZnO

F.D. Auret<sup>1</sup>, S.A. Godwin<sup>1</sup>, M. Hossain<sup>2</sup>, M.J. Legodi<sup>1</sup>, H.A. van Laarhoven<sup>1</sup> and D.C. Look<sup>3</sup>

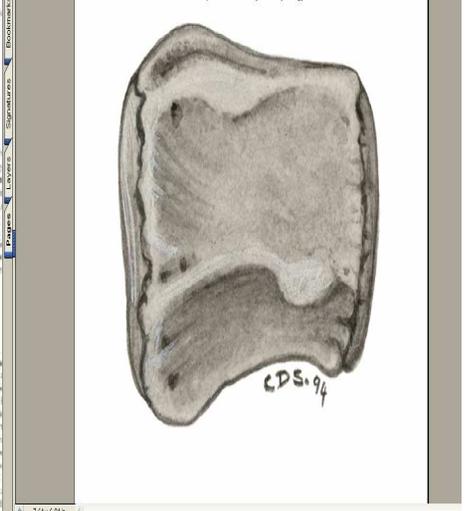
<sup>1</sup> Physics Department, University of Pretoria, Pretoria 0002, South Africa  
<sup>2</sup> Nanoscale Research Center, Wright State University, Dayton, OH, 45436, USA, and  
<sup>3</sup> Materials and Manufacturing Directorate, Air Force Research Laboratory, WPAFB, OH, 45433, USA

Received 8 June 2010  
Published 20 September 2010  
Online at stacks.iop.org/JPhys:CSM0/3/0989

**Abstract**  
We report on the electrical and defect characterization of an Schottky diode formed on single-crystal ZnO, before and after irradiating with high-energy (1.25 MeV) protons. Prior to bombardment we observed that several defect traps (E1–E4), with energies between 0.10 and 0.87 eV below the conduction band, are present in the ZnO. High-energy proton bombardment introduces two electron traps (Ep1) and Ep2), with extremely low introduction rates of 2.4 and 1.9 cm<sup>-2</sup>, respectively. Schottky barrier properties such as reverse leakage current deteriorated from  $1 \times 10^{-10}$  A for an unannealed diode to  $1 \times 10^{-7}$  A after bombardment with a dose of  $4.2 \times 10^{18}$  cm<sup>-2</sup> protons. Compared to GaN we found that ZnO is remarkably resistant to high-energy proton bombardment.

### 1. Introduction

ZnO, a wurtzitic semiconductor material with a high band gap, is presently used in its diverse products. Its main application is phosphors, paints, piezoelectric transducers, varistors and transparent conducting films, the latter being important for the photovoltaic industry. However, from a recent review, where the properties of ZnO are summarized [1], it is clear that ZnO can be used for several other, more sophisticated, electro-optical applications. On the fact that ZnO has an experimental direct band gap of 3.8 eV, it can play an important role in realizing blue and ultra-violet (UV) light emitting devices, such as light emitting diodes and lasers, as well as day-light-blind UV detectors, as in the case for GaN with a similar band gap. Furthermore, the large band gap of ZnO renders it suitable for the fabrication of a cell, catalysis and as a substrate or buffer layer for the group III nitride based devices. In some applications, these devices often have to operate at elevated temperatures, typically above 200 °C, in harsh radiation conditions comprising energetic particles. Further pro-





## CSIR context



- The CSIR turned 60 in October 2005. This event marks major changes in our organisation.
- The CSIR was given the responsibility to drive South Africa's **Open Source** initiative.
- The CSIR previously sold its publications but towards the end of 2005 that activity was stopped and the organisation started investigating **Open Access**.
- In October 2006 the Information Services took on the responsibility to establish an institutional repository for the organisation.
- UP agreed to help us learn and we thought we would have the repository set up in January 2007
- Fortunately we have been managing the CSIR's internal Technology Outputs database (which was established in the early 80's) since 2004 and did not have to start from scratch to find content.





## Fast tracking ... learning from UP



- Easy we thought: They (UP) had done the research, had gained the experience and were able and willing to train us ...
- For example we needed to know:
  - Which is the best repository software to use? **UP shared their research results.**
  - What items and under what circumstances may you place items in the repository? **UP directed us to SHERPA.**
  - Why use handles? **Introduction to Handles.net.**
  - Even ... why Dublin Core? **Notes on standards.**
  - What formats? **Oops ... they showed us the help file!**
  - What documentation needed to be prepared? **Gave us their templates.**
- The next set of questions related to the technology
  - Is vanilla version sufficient? **Facilitated meeting with IT.**
  - What is the server set-up? **Our techie was introduced to their techie.**
  - How does one manage security? **More techie meetings!**
  - How does the metadata-data 'thing' work? **Demonstration!**
- However, even with the help of our friends and with the best intentions in mind, the January launch did not become a reality.



## Experimenting ... would it be worth your while?



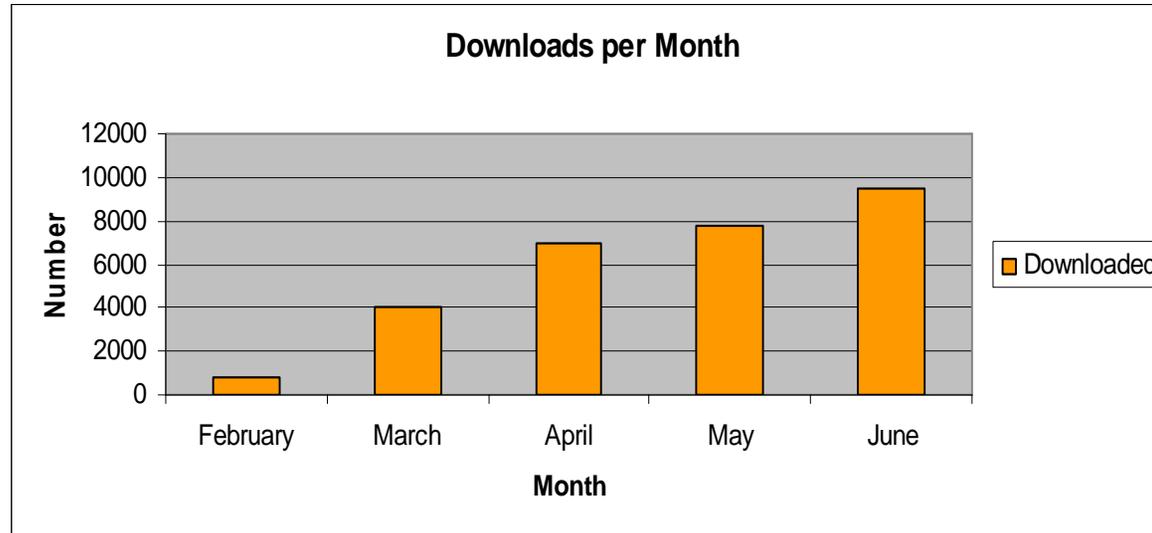
- We decided to slow down and to experiment ... to establish if a repository would be suitable for our environment.
  - UP allowed us to build a collection inside their repository – so that we could gain experience in using the DSpace software. A set of 450 documents was identified and uploaded.
  - Simultaneously our IT department revived the CSIR's eShop software ... to test the repository concept in existing proprietary software with the same set of documents.



- We intended to compare effort.
- We did not communicate the fact that we had added open access content to the CSIR site ... BUT ... we had forgotten that the CSIR-site is indexed by Google.
- ... then IT phoned to ask if we would like to see the activity associated with the 'repository'.



## Statistics



- By end April just < 6 000 documents had been downloaded.
- Group Manager R&D Outcomes was informed and we took a serious look at the way in which we made the documents accessible. The additional default features of repository software was no longer an option.
- *By end June this figure has become >28 000 documents.*



## Favourites

### February

- 22 Catastrophic failure of a raise boring machine during underground reaming operations
- 16 Impacts of air pollutants on vegetation in developing countries
- 15 Some failure analyses of South African Air Force aircraft engine and airframe components

### March

- 50 Physio-climatic classification of South Africa's woodland biome
- 43 Risk assessment as a management tool used to assess the effect of pesticide use in an irrigation system, situated in a semi-desert region
- 42 I think condoms are good but, aai, I hate those things: condom use among adolescents and young people in a Southern African township

### April

- 93 Identification of atractyloside by LC-ESI-MS in alleged herbal poisoning
- 59 Physio-climatic classification of South Africa's woodland biome
- 46 Costs and benefits of biological control of invasive alien plants: case studies from South Africa
- 46 Detective quantum efficiency of medical x-ray image intensifiers
- 46 Failure of rural schemes in South Africa to provide potable water

### May

- 63 Costs and benefits of biological control of invasive alien plants: case studies from South Africa
- 65 Assessing the effects of human-induced land degradation in the former homelands of northern South Africa with a 1 km AVHRR NDVI time-series
- 67 Risk assessment as a management tool used to assess the effect of pesticide use in an irrigation system, situated in a semi-desert region

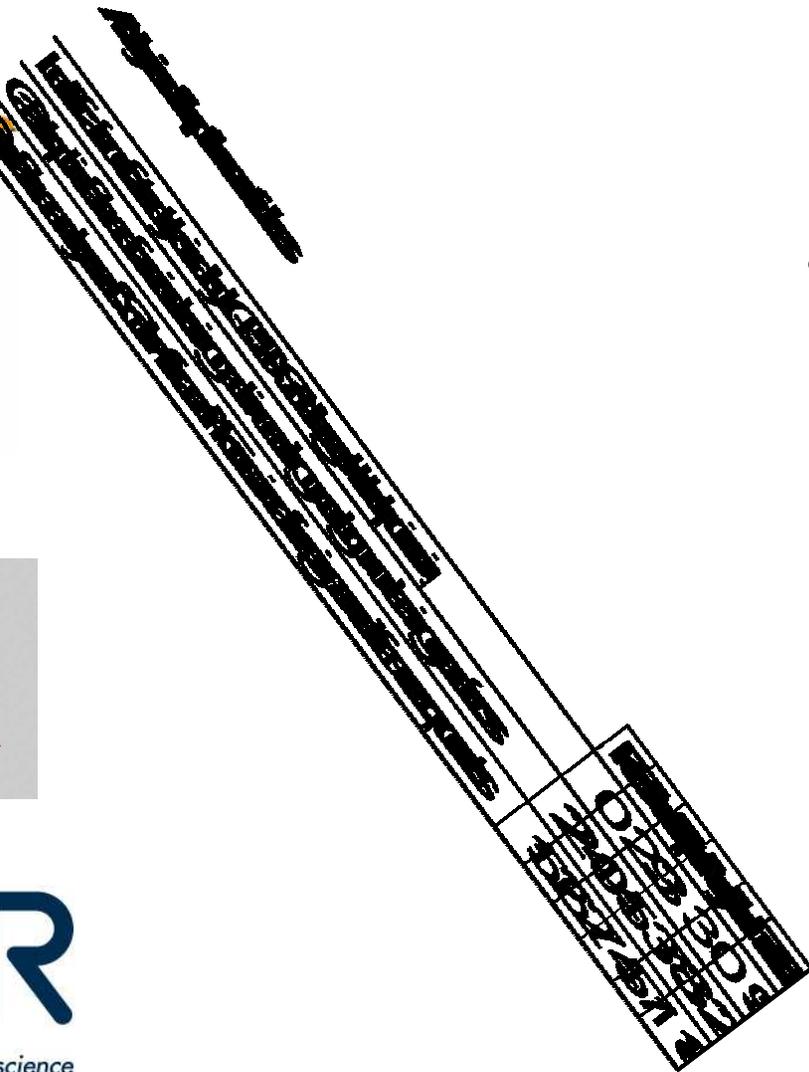
### June

- 83 Catastrophic failure of a raise boring machine during underground reaming operations
- 71 Some failure analyses of South African Air Force aircraft engine and airframe components
- 54 Nitrate removal with reverse osmosis in a rural area in South Africa
- 54 Checking anion-cation charge balance of water quality analyses: Limitations of the traditional method for non-potable waters





## Tracking individual items



- Realisation: the organisation could start tracking the impact of its research output much quicker and identify key expertise sooner.
- There was no turning back and as a team (together with IT and our Communications staff, and under the leadership of our Group Manager R&D Outcomes) we started redesigning the CSIR Internet site in such a way that the repository became an integral part of our knowledge dissemination strategy.





# Giraffe leap: our newly launched repository

The Council for Scientific and Industrial Research (CSIR) in South Africa is one of the leading scientific and technology research, development and implementation organisations in Africa. It undertakes directed research and development for socio-economic growth.

[Home](#) | [About us](#) | [Research & development](#) | [Technology transfer](#) | [Join the CSIR](#) | [Contact](#)

## Research areas

- Biosciences
- Built environment
- Defence, peace, safety & security
- Information & communications technology
- Laser technology
- Materials science & manufacturing
- Mobile intelligent autonomous systems
- Nanotechnology
- Natural resources & the environment
- Space technology
- Synthetic biology
- Publications**
- Tests and services**

[home](#) » [www](#) » CSIR homepage

## Recent Studies



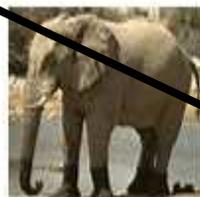
### Novel robot demonstrator for industrial use

A novel robot, using parallel technology and designed according to the mechatronics engineering approach, is being developed by the CSIR. The technology demonstrator is intended for industrial use as either a welding or spray-painting robot.



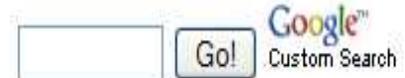
### Natural fibre composites to be used in construction

Researchers are investigating the application of natural fibre composites (NFCs) in construction, with the aim of producing construction materials that are non-harmful, recyclable and made from renewable materials.



### 130 000-year-old fossil elephant found near Durban, South-Africa

A modern African elephant tusk was discovered within a late Pleistocene aeolianite at the base of an intertidal rock pool at Reunion Rocks, south of Durban. Dating indicates that the tusk is older than 112 kyr BP.



## News



[CSIR to host tracking and mapping workshops in August](#)



[CSIR investigates use of natural fibre composites in construction](#)

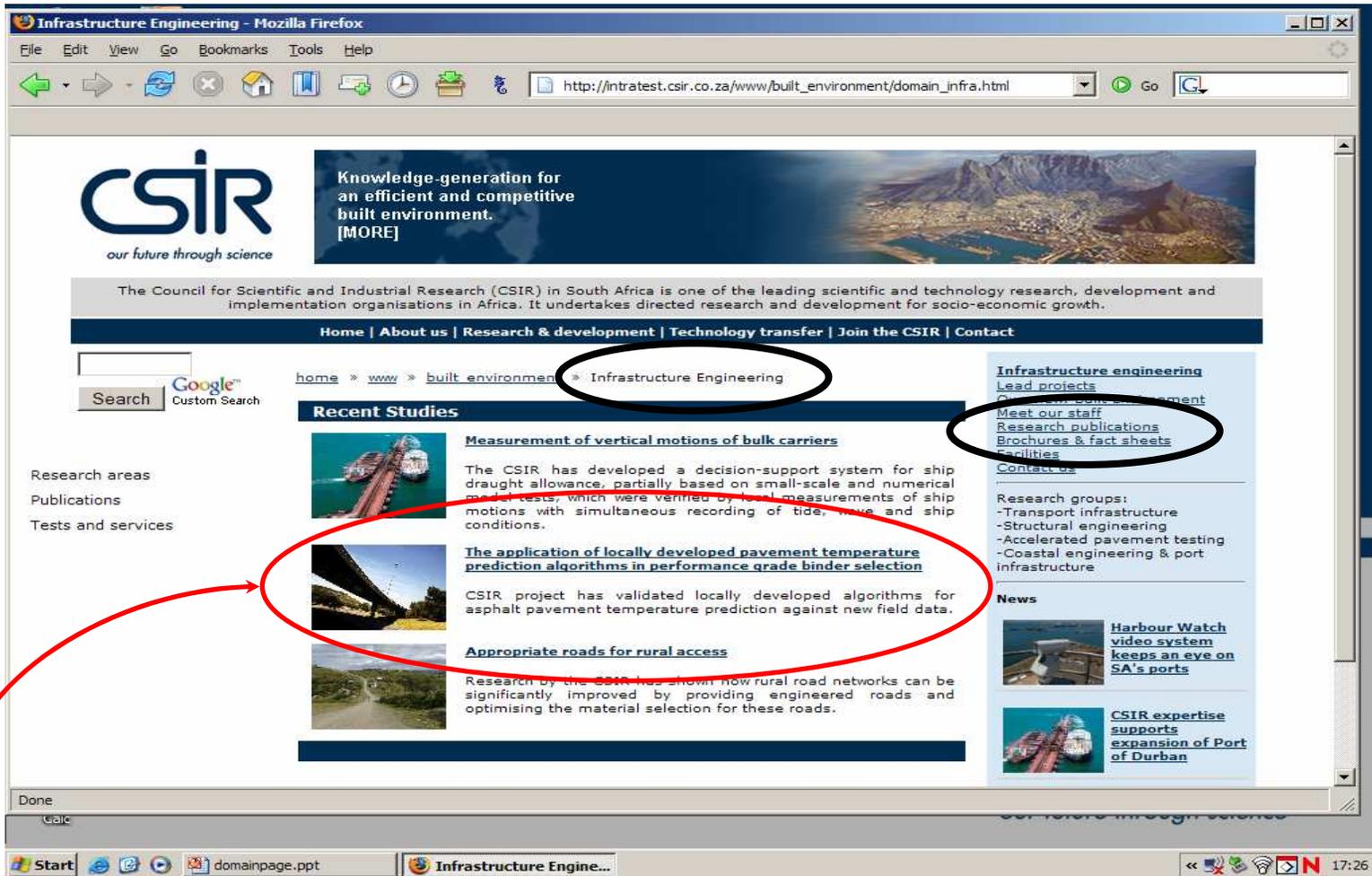


[Research network for advanced robotics and mechatronics](#)

[More news...](#)

The same in ResearchSpace

# Also accessible from competence area



Joint effort ... drawing upon the knowledge and expertise of our communications practitioners when new publications are posted in ResearchSpace

# ResearchSpace: access to a subset of our explicit knowledge



The screenshot shows a Mozilla Firefox browser window displaying the CSIR Research Space website. The address bar shows the URL <http://researchspace.csir.co.za/dspace/>. The website header features the CSIR logo and the tagline "our future through science". The main content area is titled "CSIR Research Space" and includes a welcome message, a description of the database, and a search box. The left sidebar contains navigation links for "Search ResearchSpace", "Browse", and "Sign on to:". The right sidebar contains the text "CSIR Research Space" and a note about content updates.

CSIR Research Space: Home - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://researchspace.csir.co.za/dspace/

Google Advanced Sc... Corporate Identity Te... Add to Connotea Connotea: mvandeve'...

copernic Search Web Search Desktop: All

**CSIR**  
our future through science

CSIR Research Space  
Open and unrestricted access  
to our research outcomes

Search ResearchSpace  
 Go  
Advanced Search  
Home

**Browse**  
Communities & Collections  
Titles  
Authors  
Subjects  
By Date

**Sign on to:**  
Receive email updates  
My ResearchSpace authorized users  
Edit Profile  
Help  
About DSpace

CSIR Research Space >

### CSIR Research Space

Welcome to CSIR Research Space.

This database provides access to some of the research outputs generated by CSIR scientists. The CSIR in South Africa is one of the leading scientific and technology research, development and implementation organisations in Africa. It undertakes directed research and development for socio-economic growth.

**Note:** The mapping from the General Science, Engineering and Technology Collection to the other communities, sub-communities and collections will be completed by 1 August 2007, in time for the official launch of **CSIR Research Space**.

By 08 August 2007, 857 papers have been uploaded, with continuous additions to the database.

Please use the **Advanced Search** option for more detailed searches.

While care has been taken with uploading metadata, there may be some errors. Please feel free to report any problems to us.

### Search

To search CSIR ResearchSpace, please enter your search terms in the box below.

Go

**CSIR Research Space**

Content on this site is updated on a continuous basis. For more information regarding the contents of CSIR Research Space, or any access problems or concerns please contact [Adele van der Merwe](#)

Done

Start D:\Myfiles\P... Novell Group... Mail To: Siph... Microsoft Po... CSIR Rese... EN copernic 17:22



# Communities & collections



The screenshot shows the CSIR Research Space website in a Mozilla Firefox browser window. The page title is "CSIR Research Space: Communities and Collections". The main content area displays a hierarchical menu structure with three main categories, each with a list of sub-items:

- Biosciences**
  - Agroprocessing and chemical technology
  - Analytical science
  - Aptamer technology
  - Bioprospecting
  - Discovery chemistry
  - Enzyme technologies
  - Microbial expression systems
  - Plant biotechnology
  - Structural biology
  - Systems biology
  - Yeast expression systems
- Built environment**
  - Architectural sciences
  - Construction
  - Infrastructure engineering
  - Infrastructure systems and operations
  - Logistics and quantitative methods
  - Rural infrastructure and services
  - Simulation modelling and spatial analysis
- Defence, peace, safety & security**
  - Aeronautic systems
  - Landward sciences
  - Optronic sensor systems
  - Radar and electronic warfare systems
  - Safety and security

Arrows from the text "Same structure as throughout the Internet site" point to these three main categories.

Same structure as throughout the Internet site





## Knowledge we are 'capturing'



- At this stage only published research articles – copyright cleared via SHERPA/Romeo.
- Selected research reports will be added next. Especially then, those reports that came about as a result of publicly funded research.
- The repository software makes provision for variety of formats. We'll therefore also be adding data sets, source code, sound, video and graphic files when these become available.





## The advantage/disadvantage of leapfrogging



- The standard time it takes to set up a successful repository is reported to be three years.
- UP, because of the knowledge they gained from the experience of others, (and with a lot of hard work) reduced that time to two years.
- The CSIR, in turn, was able to create its repository within 12 months.
- We have now added our own learning and are a point where we can start to reciprocate.
- On the flip side: do not expect to have the same *in depth* understanding as the original investigator.
- Much patience/time is needed! Transfer of knowledge needs to be planned and managed or it could become all consuming!





## Tips, trick & traps



- Commitment from organisation leadership is important ... to ensure that there is focus as well as for the investment that will ensure sustainability.
- The library is now also a “publisher” of information and knowledge and not only a user and conduit.
- If you allow IT infrastructure and/or software to be stumbling blocks ... they will be. The repository software is not the issue. The repository is.
- It is important to have the courage to experiment.
- It is lonely and frustrating to play by yourself. Make others aware that you would like to play and they'll gladly help you leapfrog.
- Unfortunately: playing leapfrog is not the same as playing catch-up. The first step is to establish trust then be prepared to learn, to experiment, to work hard and to have fun.





## Lessons & tips cont



- Do not plan the perfect repository. Get the basics under control and start experimenting!
- The Internet and especially open source and open access accelerate learning. Tips and techniques are available from a wide variety of sources!
- Personal knowledge transfer and personal networks remain key to ensuring successful learning partnerships.
- Be prepared to invest time and resources. CSIR had to invest in a new server and employ extra hands before we could get going and if that is what it takes to get your repository on-line make the commitment.
- Giraffes play ... they do not travel barefoot over the Drakensberg!
- Africa's knowledge needs to be set free ... just imagine all those giraffes playing leapfrog!
- Be warned: Once the organisation KNOWS the benefits there is no back tracking ... all systems go!



**Questions?**



## Sources to consult ... as a start

- Beaudoin, Patsy and Margret Branchofsky, "MIT's DSpace Experience: A Case Study." (2004) <http://www.dspace.org/implement/case-study.pdf>
- Budapest Open Access Initiative, *A Guide to OAI-compliant Institutional Repository Systems* (2004). <http://www.soros.org/openaccess/software>
- Campbell, Lorna M., Kerry Blinco, Jon Mason (2004) *Repository Management and Implementation*. Available: [http://www.jisc.ac.uk/uploaded\\_documents/Altilab04-repositories.pdf](http://www.jisc.ac.uk/uploaded_documents/Altilab04-repositories.pdf)
- Centre for research Libraries: <http://www.crl.edu>
- Chan, L. (2004), "Supporting and Enhancing Scholarship in the Digital Age: The Role of Open Access Repositories", *Canadian Journal of Communications*, 29 (3-4), 277-300.
- Clifford A. Lynch, "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" *ARL*, no. 226 (February 2003): 1-7.
- Crow, Raym. *The Case for Institutional Repositories: A SPARC Position Paper*. SPARC: Scholarly Publishing & Academic Resources Coalition (2002). [http://www.arl.org/sparc/IR/IR\\_Final\\_Release\\_102.pdf](http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf)
- DSpace Project Implementing DSpace section of DSpace.org website: <http://dspace.org/implement/index.html>
- DSpace: Free, open source software jointly developed by MIT and Hewlett Packard Labs. Download at <http://sourceforge.net/projects/dspace>
- Mary R. Barton (2004) *Creating an Institutional Repository: LEADIRS Workbook*
- Pathfinder Report on Web Based Repositories available: <http://www.palsgroup.org.uk>
- SHERPA: <http://www.sherpa.ac.uk/documents>
- SPARC Europe: <http://www.sparceurope.org/index.html>