Content Management models: an overview

Patricia Liebetrau

Digital Resource Management from Data to Knowledge 21 – 23 July 2009, Durban, South Africa



Introduction

- Increasing digitisation efforts
 - Increasing data
 - Dispersed digital projects / collections
 - Digital libraries
 - Cross institutional subject repositories
- New models for digital content management
- Interoperable digital libraries
- Networked infrastructure harvesting protocols



Considerations ... 1

- Proprietary software?
 - Licence fees
 - Customer support
 - Low IT expertise required
 - Upgrades extra cost?
 - Warranty/ maintenance?
- Open source software?
 - Free downloadable source software
 - Customisation hidden costs
 - IT expertise good user documentation
 - Community support
 - Open standards
 - Software tools (parsing/editing/transforming/validating)

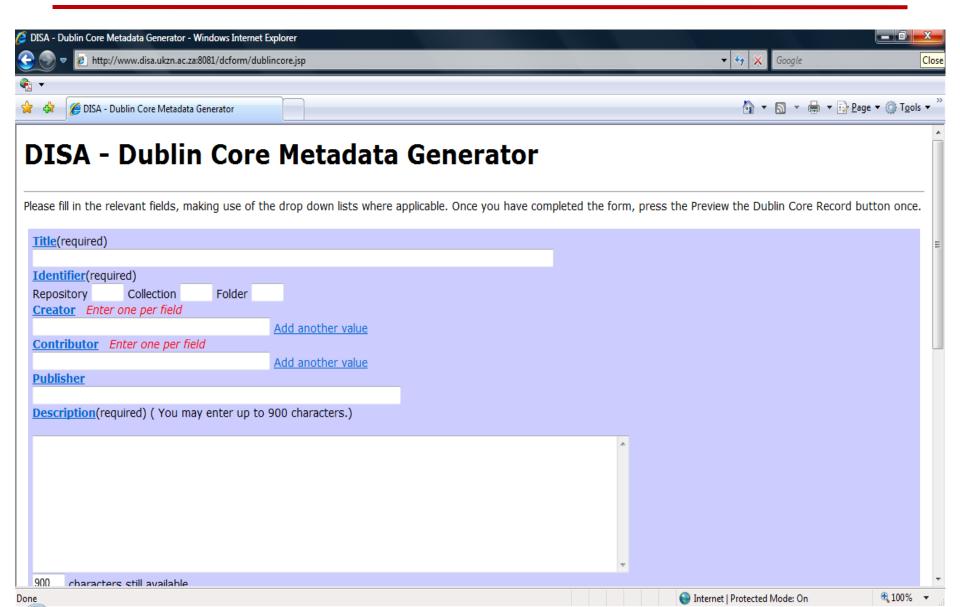


Considerations ... 2

- Content Management software packages
 - Greenstone
 - DSpace
 - EPrints
 - Fedora
 - Joomla
 - Drupal
 - Many others
- Modular DIY
 - Open source database Postgress, MYSQL
 - Folder structures
 - Online input forms (DC and METS)



Online input form



Considerations ... 3

- Open access
 - Metadata and full text available without restrictions
 - No downloading charges
 - Social networking services delicious, twitter,
 - Openness of attitude
 - Core functions of HEI
- Closed access and subscription based
 - Restricted access to the site
 - Restricted access to full text
 - Abstracts available, full text access requires subscription
 - Annual subscriptions individual or institution



International standards

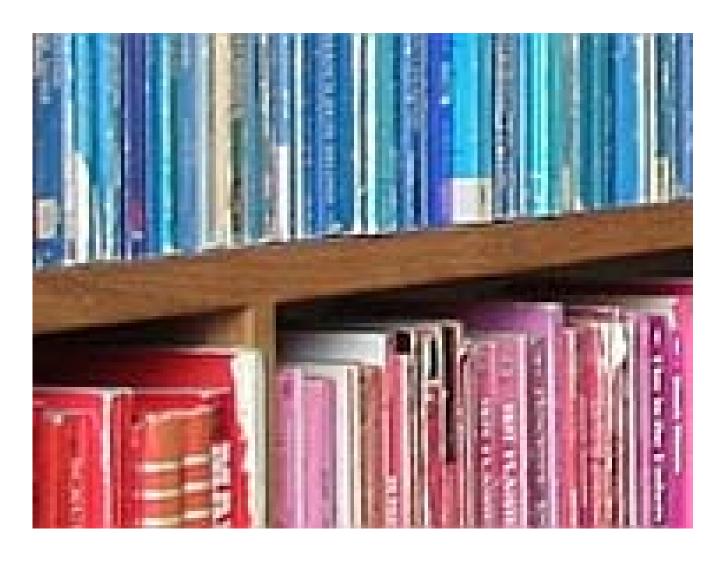
- XML is the *lingua franca* of the web
- XML de facto standard for metadata
- Metadata element sets / Schemes



Community standards

- Archival Community
 - EAD?
- Heritage Community
- Arts and culture
- Literature
 - TEI?
- Institutional Repositories
 - Dspace
 - Eprints
- Small digital projects especially non-web access and distribution
 - Greenstone





We are doing things differently

African context — issues to consider

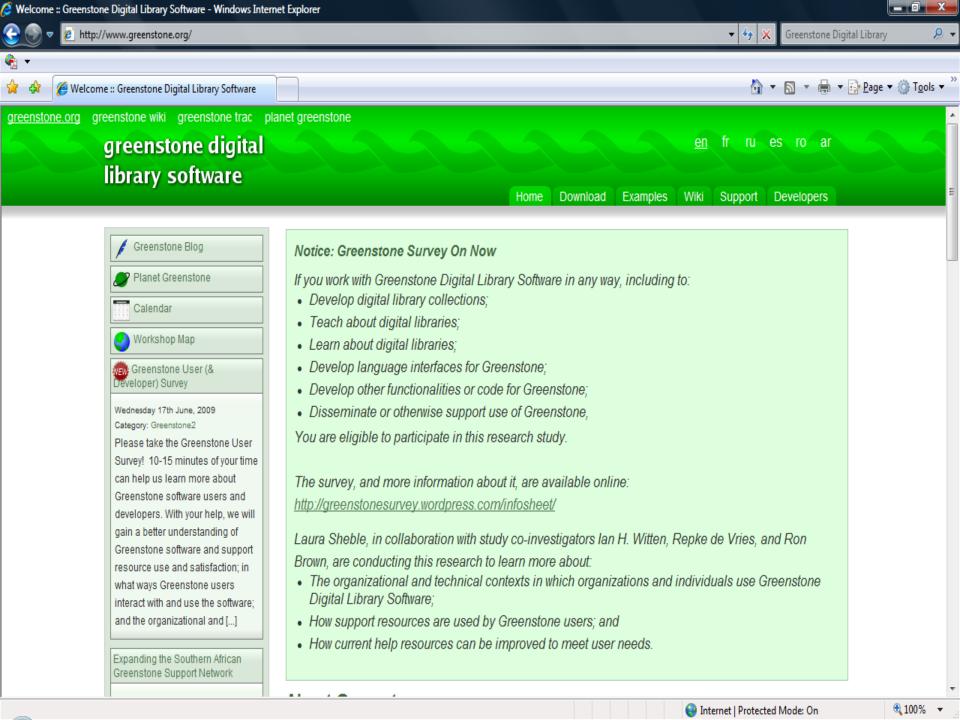
- Available resources
- Expertise especially IT
- Open source software
- Extensibility
- Interoperability
- Robustness and stability
- Access (web or other medium such as DVD and CD)
- Sustainability
- Bandwidth



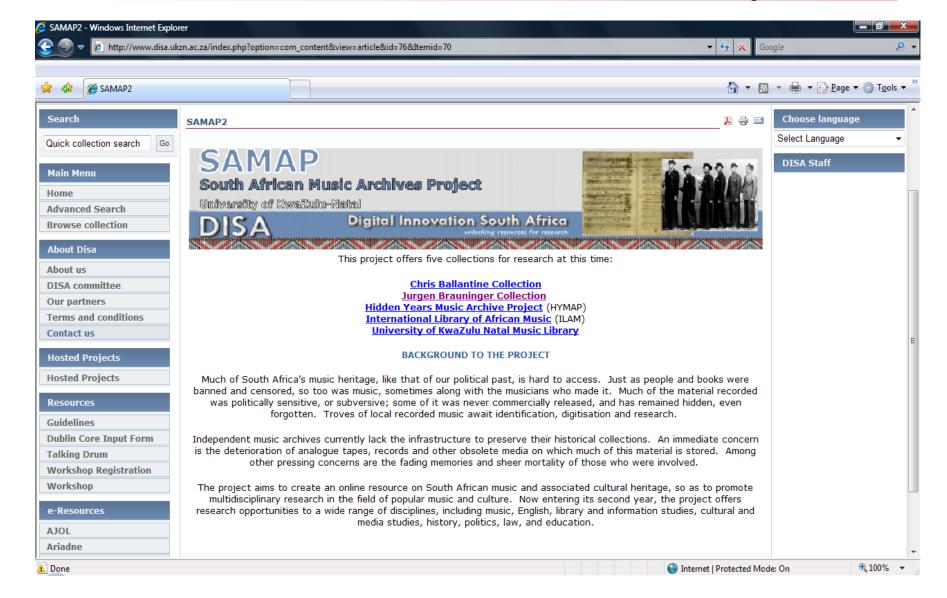
Greenstone

- A suite of software
- Open source (community development)
- Build digital collection/library/ project
- Management and access (web/ CD/DVD)
- Add metadata
- Customise for your requirements (branding)
- Search and browse functions

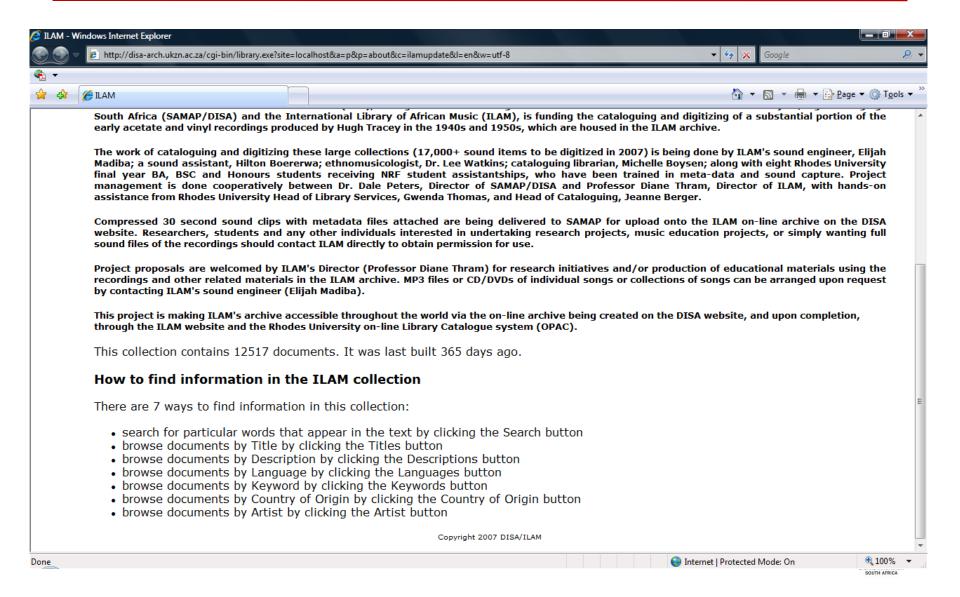




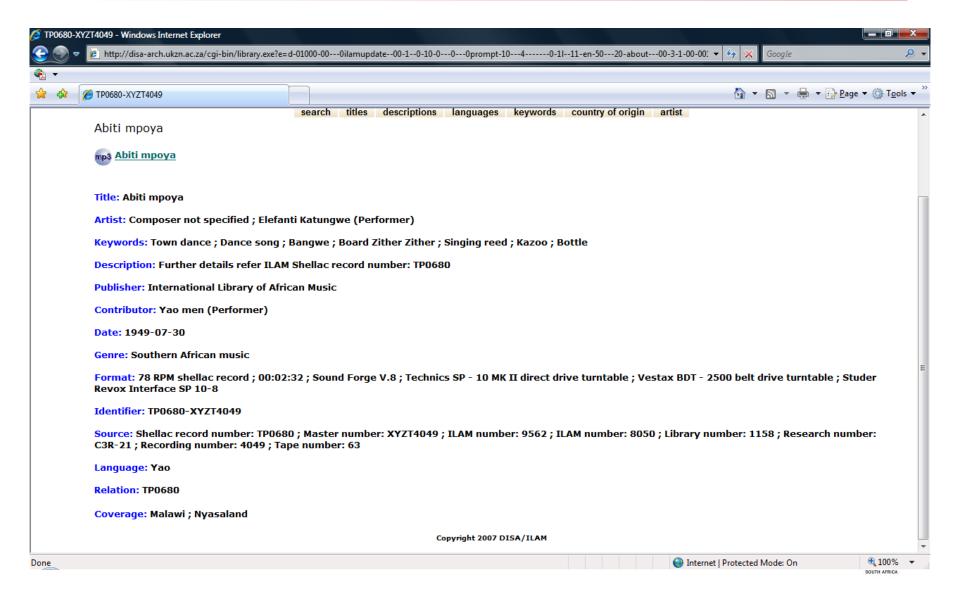
SAMAP South African Music Archive Project



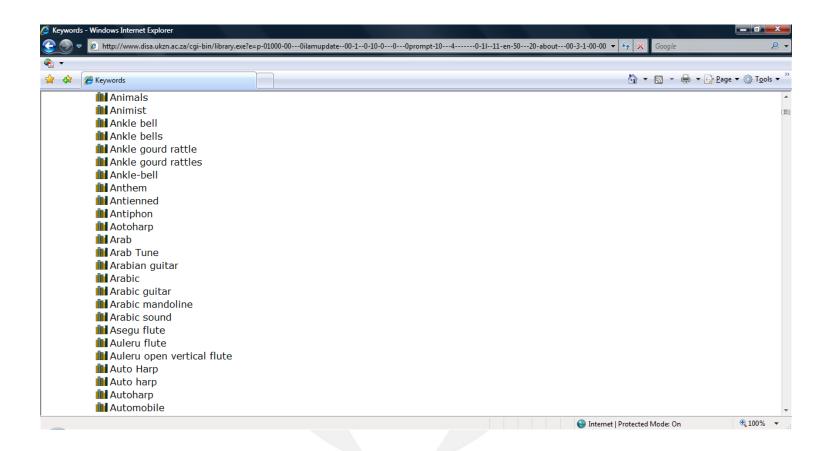
Search and browse functions



Metadata display

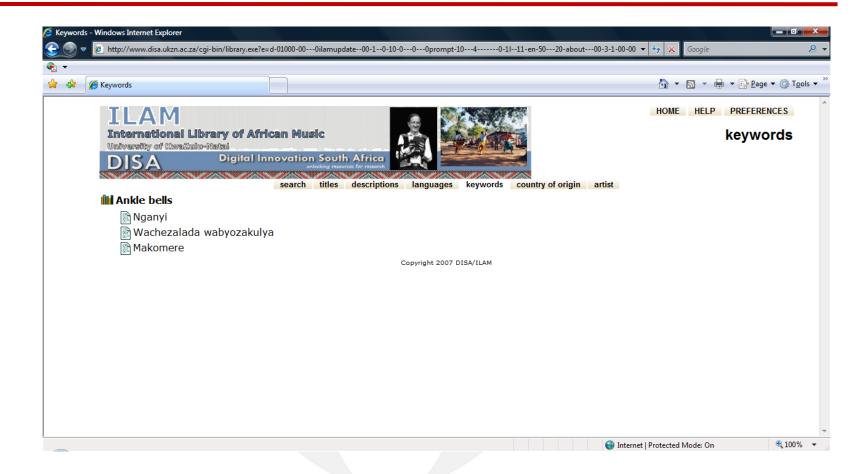


Subject keywords



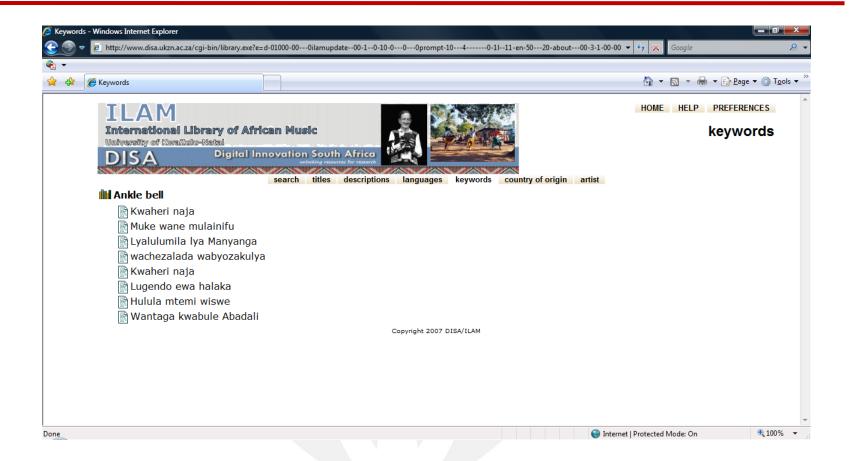


Ankle bells





Ankle bell





Overview of Greenstone

- Positive:
 - Ease of startup
 - Community support
 - CDs/DVDs for distribution, exchange, sales, promotion



Overview of Greenstone

Positive:

- Ease of startup
- Community support
- CDs/DVDs for distribution, exchange, sales, promotion

Drawbacks:

- Extensibilty and robustness limitations
- Customisation time consuming
- Large collections → long builds
- Interoperability limitations
- Precise searching → unsatisfactory search results



Institutional Repository development

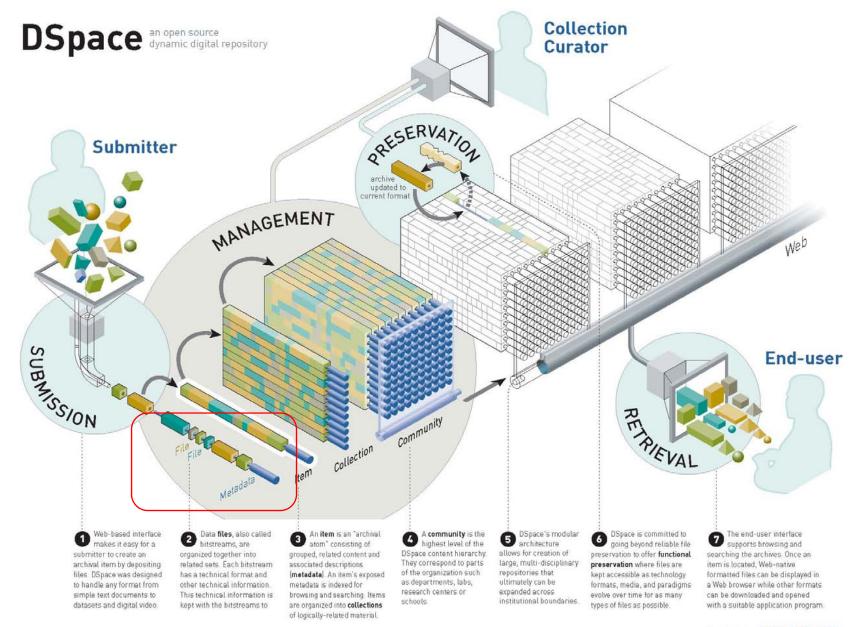
- Research output
 - Students
 - Theses
 - Staff
 - Pre-prints and academic research papers
 - Research data sets
 - Conference papers
 - Coursework
 - Examination papers
- ▶ In Africa 24+
- ▶ In SA 16+
- Most popular content management software is DSpace



What is D-Space?

- Open source software
- Suite specifically designed for Institutional Repositories
- Development and management
- Based on international standards
- Interoperability





Subject repositories

- Responsibility of libraries, archives and museums, as aggregators of research resources, to organise diverse information for online resource discovery.
- Subject repositories may span across several collaborating institutions
- ▶ Potential component of developing an digital repository model that offers an alternative structure of scholarly communication in South Africa.

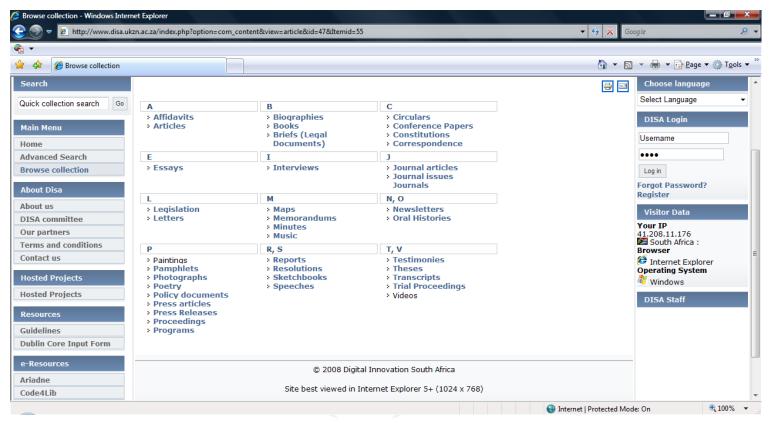


Digital subject repository

"Digital archives or digital libraries which collect and manage resources relating specifically to one or more subjects".

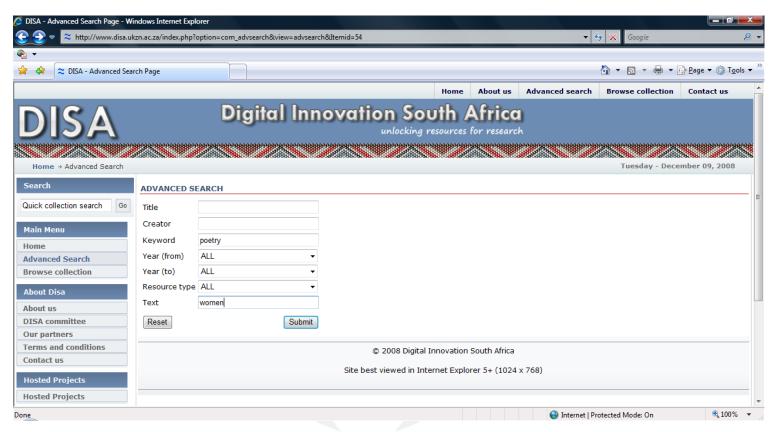


DISA browse according to type of resource



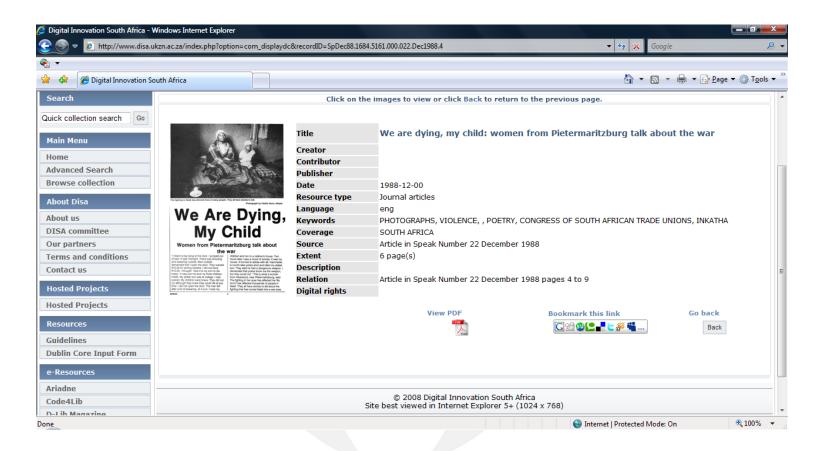


Advanced search





Metadata and full text





Digital content management

- MySQL
- Metadata Input online input form
- Folder structure
- Joomla web content management
- Fedora content management



Fedora

- Open source
- IT expertise for customisation
- Community user support
- Collaboration initiative between D-Space and Fedora
- Interoperability
- Lightweight customisation development

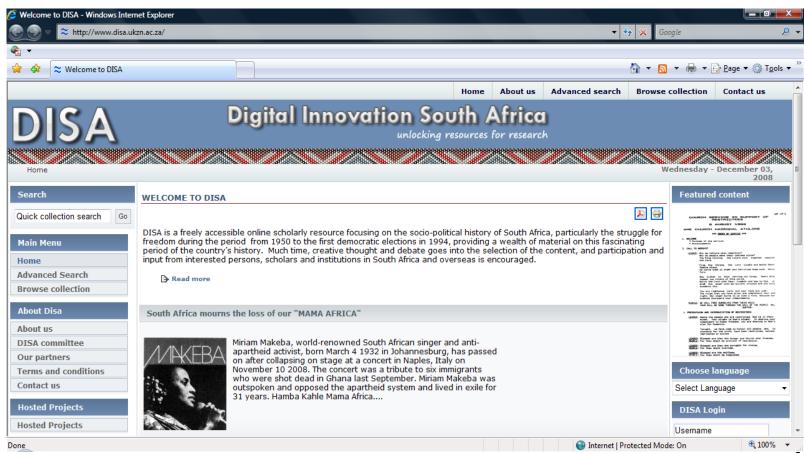


Web site management software

- Open source
 - Joomla
 - Drupal
- Templates
 - Home page design
 - Colours
 - Liquid screen
 - Customisable for branding and logos
 - Authentication for confidentiality
- Modules
 - Upload documents
 - RSS feeds
 - User



DISA: Digital Innovation South Africa



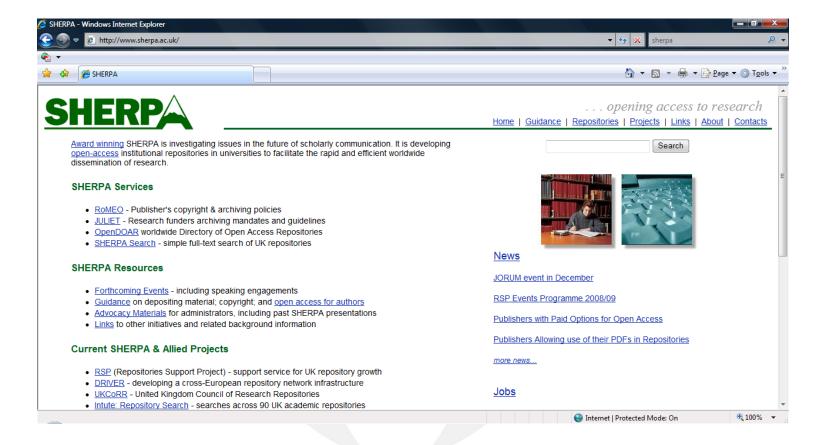


Networks

- Development of cross-repository infrastructures
- Metadata interoperability
- Standards

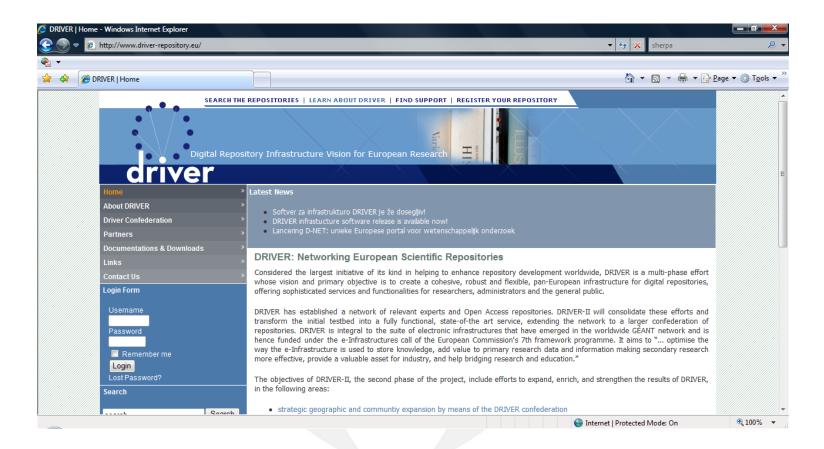


SHERPA



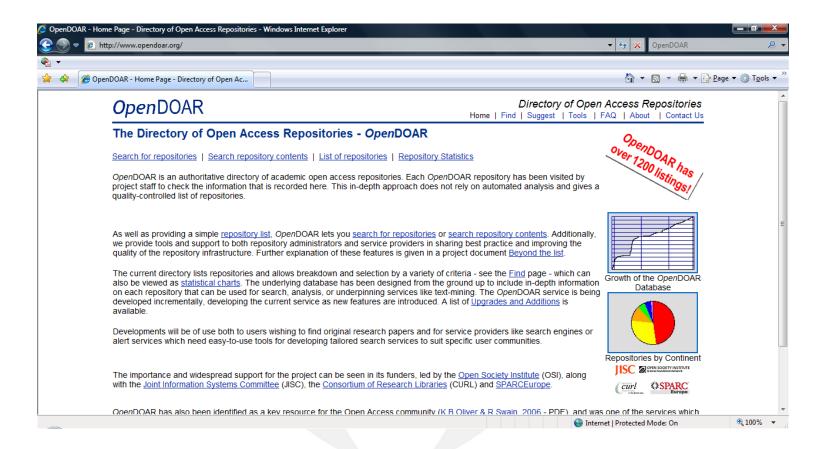


DRIVER





OpenDOAR





Implementation framework

- Administrative functions
 - Project management
 - Digital rights management
 - Staff
- Technical issues
 - Software platform
 - Data management
- Search services
 - Ingest
 - Search
 - Retrieval
- Content management
 - Ingest
 - Backups
 - Persistent archive



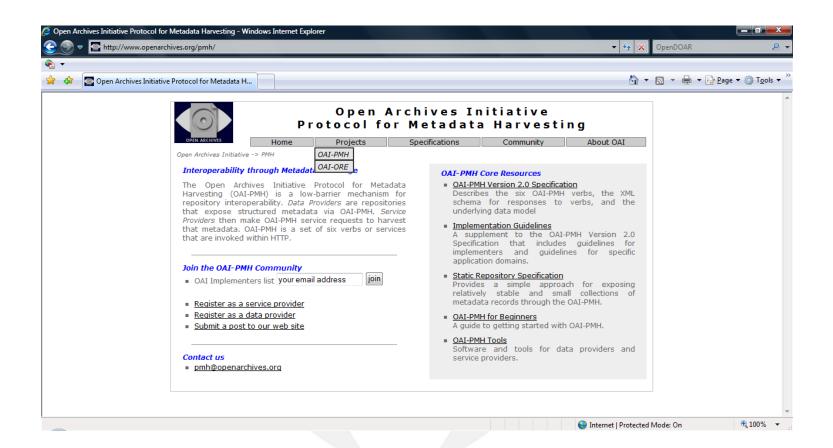


OAI-PMH Harvesting Protocol

- Open Archives Initiative : Protocol for Metadata Harvesting
- Standard for exchange of metadata
- Data providers employ the OAI-PMH to expose structured data, metadata, in various forms.
- Service providers use the OAI-PMH to harvest the metadata from data providers
- Automatically process it and add value in the form of own services.
- Measure of success in establishment of communities and/or research projects that collectively agree to adopt the protocol as a basis for information federation.
- HTTP protocol: <oai> GetRecord



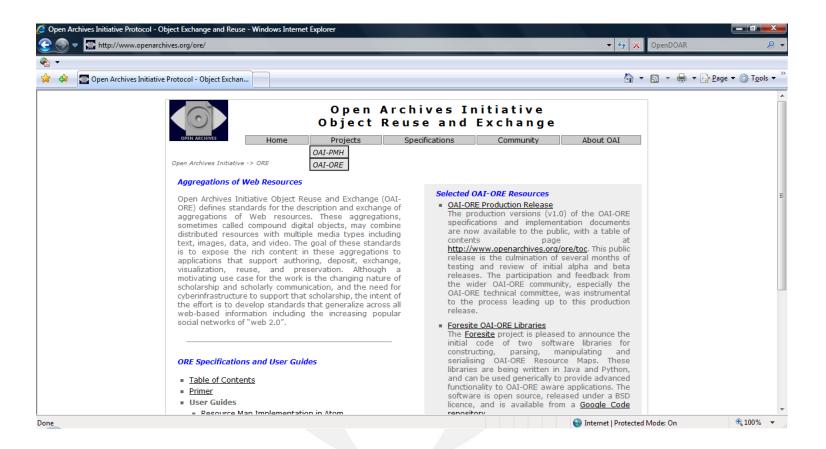
OAI-PMH







OAI-ORE





Conclusions

- Select a CMS according to your needs
- Consider IT requirements
- Consider user community requirements
- Consider growth of your collection/repository/digital library
- Build for repository harvesting services
- Ensure interoperability
- Global repository networks



Contact details

Patricia Liebetrau, Project Manager, DISA

► Email: <u>liebetraup@ukzn.ac.za</u>

URL: http://www.disa.ukzn.ac.za

Reference:

- ▶ The future of repositories? Patterns for (Cross-)Repository Architectures.
- http://www.dlib.org/dlib/november08/aschenbrenner/11aschenbrenner.html
- This presentation is made available under a <u>Creative Commons Attribution 2.0 South Africa</u> license.

