

Applying Open Storage to Institutional Repositories



<http://www.preserv.org.uk>

David Tarrant
University of Southampton
dct05r@ecs.soton.ac.uk

It helps to be open...

IRs are largely characterised by “Openness”

Open Source Software

Open Access (Initiative)

...

Open Storage

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries

About Preserv...

How can we facilitate effective
preservation survives for
digital repositories?

Preserv2 is a Project, not a service provider.

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries



Independence is also good...

No *single* vested interest in...

Repository Software

EPrints and Fedora are represented in the project

Storage

Working with Sun among others via the CRIG

Preservation Tools

TNAs Pronom Droid, Some connections with Planets

Preservation Service Providers

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries





WWW

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries



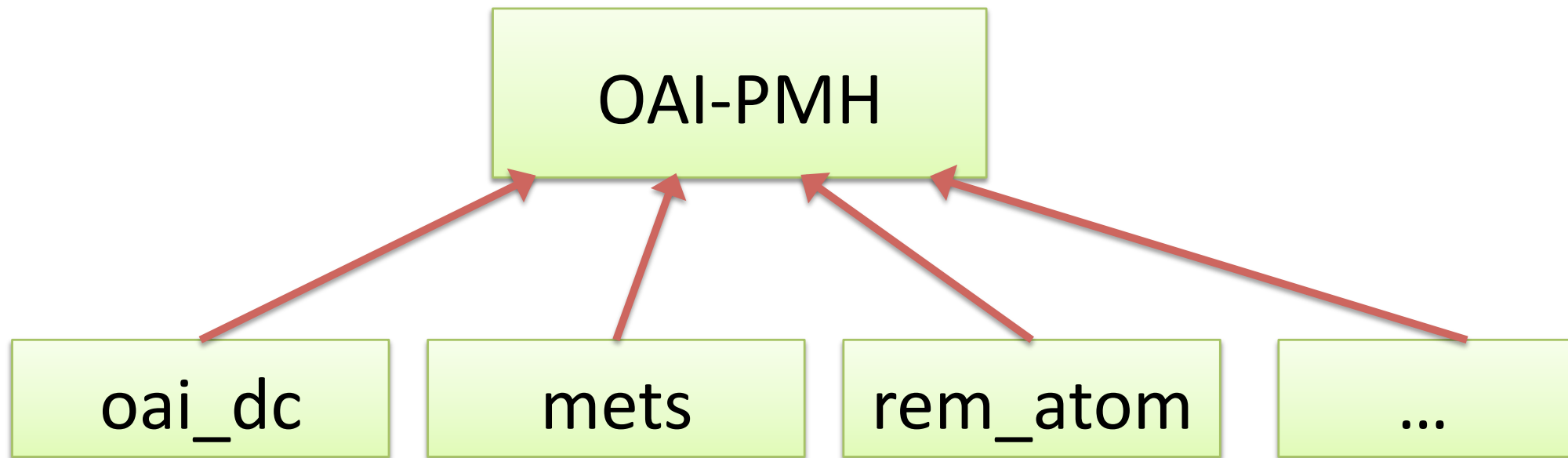
Preserv^{.org.uk}
Repository Preservation and
Interoperability

Taking the Bull by its Horns

- We can focus on building & testing an infrastructure that connects up components.
- The aim is to serve repository managers not by answering their preservation questions one at a time or per object.
- All approach, manual curation does not scale!

Interoperability

- OAI is one enabler to Interoperability



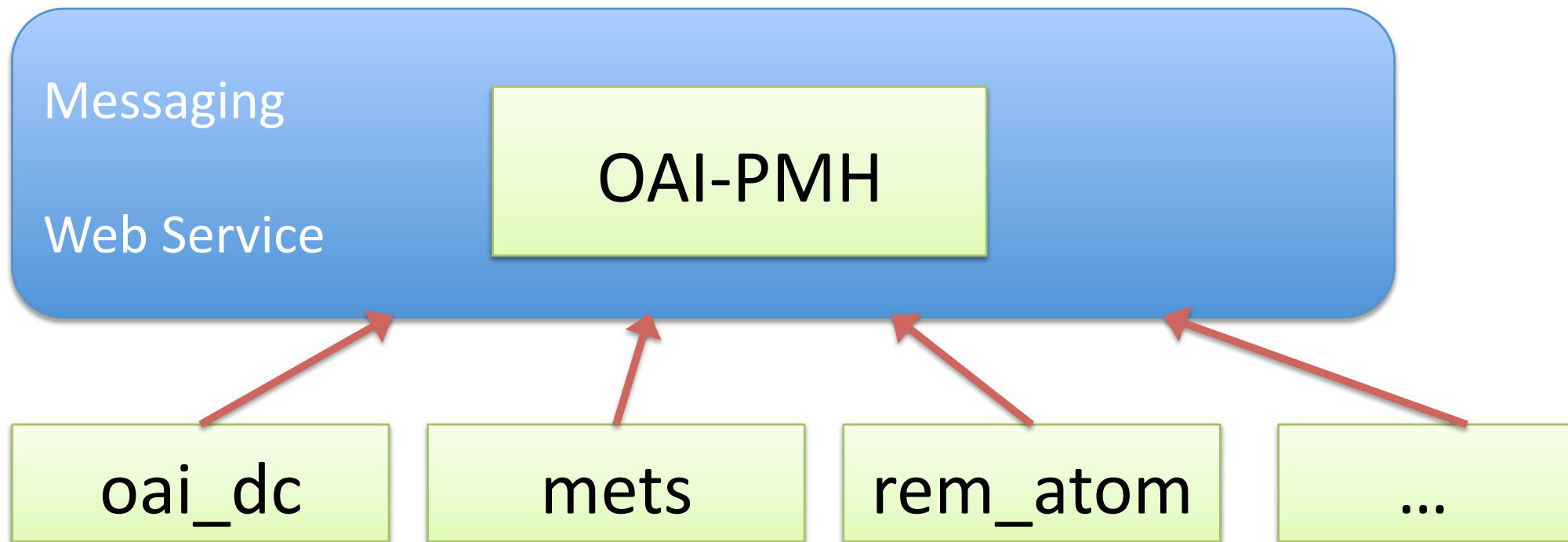
What OAI did for Repositories

Provided a standardised way to share information pertaining to Objects in a Repository.

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries

Interoperability

- OAI is one enabler to Interoperability



ORE

OAI-ORE Serialisation
REM_ATOM / REM_RDF

REM
Description + Aggregation

DC

FOXML

METS

PDF

EPXML

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries

What ORE does for Repositories

Reveals the first class objects/resources in the repository.

More than just the metadata is **guaranteed** to be described in a Resource Map (REM)

Mining with ORE



OAI-ORE EPrints & Fedora

OR08 Publications

Conference Home | Repository Home | About | Browse by Session | Browse by People | Advanced Search

Welcome to OR08 Publications

This is a repository used to manage publications for the **Open Repositories 2008** conference. All eprint URLs are persistent and will be redirected to other repositories or services as required in the future. Hint: To download many items at once, [perform a search](#) and export the results in "Zip" format.

- Main Conference**
 - Keynote
 - Repository Developer Challenge Finalists
 - Tuesday 18th April
 - 1 - Web 2.0
 - 2a - Social Networking
 - 2b - Sustainability (a)
 - 3 - Interoperability
 - Birds of a Feather (Tues)
 - Posters
 - Wednesday 2nd April
 - 4a - National Perspectives
 - 4b - Scientific Repositories (a)
 - 5a - Legal
 - 5b - Scientific Repositories (b)
 - 6a - Sustainability (b)
 - 6b - Models, Architectures & Frameworks
 - 7 - Usage
 - Birds of a Feather (Wed)
 - Repository Managers
 - Object Reuse and Exchange (ORE) Open Day
- User Groups**
 - ISpace User Group
 - Case Studies
 - OSpace 1.5 Launch and beyond
 - Engagement/Interaction
 - Getting to grips with 1.5
 - Interoperability
 - EPrints User Group
 - EPrints 3.1 and EPrints Future
 - EPrints Training and Support
 - Repository Analytics
 - Research Assessment Experience
 - Fedora User Group
 - Architecture
 - Case Studies
 - Datasets
 - Fedora EPrints
 - Proof Ends
 - Preservation and management
 - Programming
 - Search
 - Semantic Technologies

OR08 Publications supports [OAI 2.0](#) with a base URL of <http://pubs.or08.eos.oxon.ox.ac.uk/cgi/vs12>

ora
OXFORD UNIVERSITY
RESEARCH ARCHIVE

Search Detailed Search

ORA Basic Item: "Job Mobility of Residents and Migrants in Urban China"

Reference:
John Knight; Linda Y. Yueh, (2003). Job Mobility of Residents and Migrants in Urban China.
Link to this archived copy: <http://eprints.ses.ox.ac.uk/5000/object/pdf/Dec1744f-e435-4176-86e9-941003a8021>

Title
Job Mobility of Residents and Migrants in Urban China ;

Creator
John Knight; Linda Y. Yueh ;

Date
2003 ;

Subject
Classification-JEL: J21 ; Classification-JEL: J60 ; Classification-JEL: J63 ; Classification-JEL: O53 ; labour mobility ; labour turnover ; layoffs ; China ;

Format
Application/pdf ;

Downloads
[Full Text](#) (as ASCII text of item)
[Journal](#) (pdf item)

Terms of Use
The copyright of this item rests with the author and/or other copyright holder(s).
[Click here for our Terms of Use](#)

RePEc-Handle: RePEc:oxf:wpaper:163
RePEc-number: 163
Original-URL: <http://www.economics.ox.ac.uk/Research/wp/pdf/paper163.pdf>
[ora1202](#)
[um.vuud/Dec1744f-e435-4176-86e9-941003a8021](http://eprints.ses.ox.ac.uk/5000/object/pdf/Dec1744f-e435-4176-86e9-941003a8021)

Preserv
Repository Preservation and Interoperability

SUN PASIG
Spring Meeting
May 27-29, 2008

JISC / CNI
Transforming the User Experience
July 10-11, 2008

MICROSOFT E-SCIENCE WORKSHOP
Dec 7-9, 2008

OR2008
Welcome to Athens, Georgia

Oxford University Research Archive

Home | About | Browse by Year | Browse by Subject

Login | Create Account

Job Mobility of Residents and Migrants in Urban China
John Knight and Linda Y. Yueh (2003) *Job Mobility of Residents and Migrants in Urban China*.

XML 1188b

Json ["document_type_name_textcalendar" not defined] 955b

Plain Text 1362b

XML 678b

PDF - Requires a PDF viewer such as [GSView](#), [Xpdf](#) or [Adobe Acrobat Reader](#) 241Kb

Plain Text 66Kb

OR2008
Third International Conference on Open Repositories

Browse: Search Detailed Search

Date [see more](#)
2008-04 - (86)

Subject [see more](#)
Posters - (50)
1 - Web 2.0 - (3)
2a - Social Networking - (3)
3 - Interoperability - (3)
4a - National Perspectives - (3)
4b - Scientific Repositories (a) - (3)
5a - Legal - (3)
5b - Scientific Repositories (b) - (3)
6a - Sustainability (b) - (3)
6b - Models, Architectures & Frameworks - (3)

Automatically derived terms [see more](#)
4 April - (72)
application/pdf - (72)
repositories - (64)
southampton - (58)
pubs - (51)
submission - (31)
united kingdom - (30)
posters - (29)
repository - (21)
open access - (13)

Type [see more](#)
Conference or Workshop Item - (86)
NonPeerReviewed - (86)

Creator [see more](#)
Carr, Leslie - (5)
Allinson, Julie - (3)
Hubbard, Dill - (3)
Awre, Chris - (2)
Coles, Simon - (2)
Kahn, Jeffrey - (2)
Kumar, Anoop - (2)
Murray-Bust, Peter - (2)
Namiki, Takao - (2)
Pepler, Sam - (2)

Format [see more](#)
application/pdf - (81)
application/vnd.ms.powerpoint - (10)
text/html - (2)

Disclaimer and Data Protection statement | Accessibility statement

[Eprints](#) [Fedora Commons](#)

Site powered by Fedora and Apache Solr. Data source and information management system is powered by EPrints.org



Which is which?

Open Storage

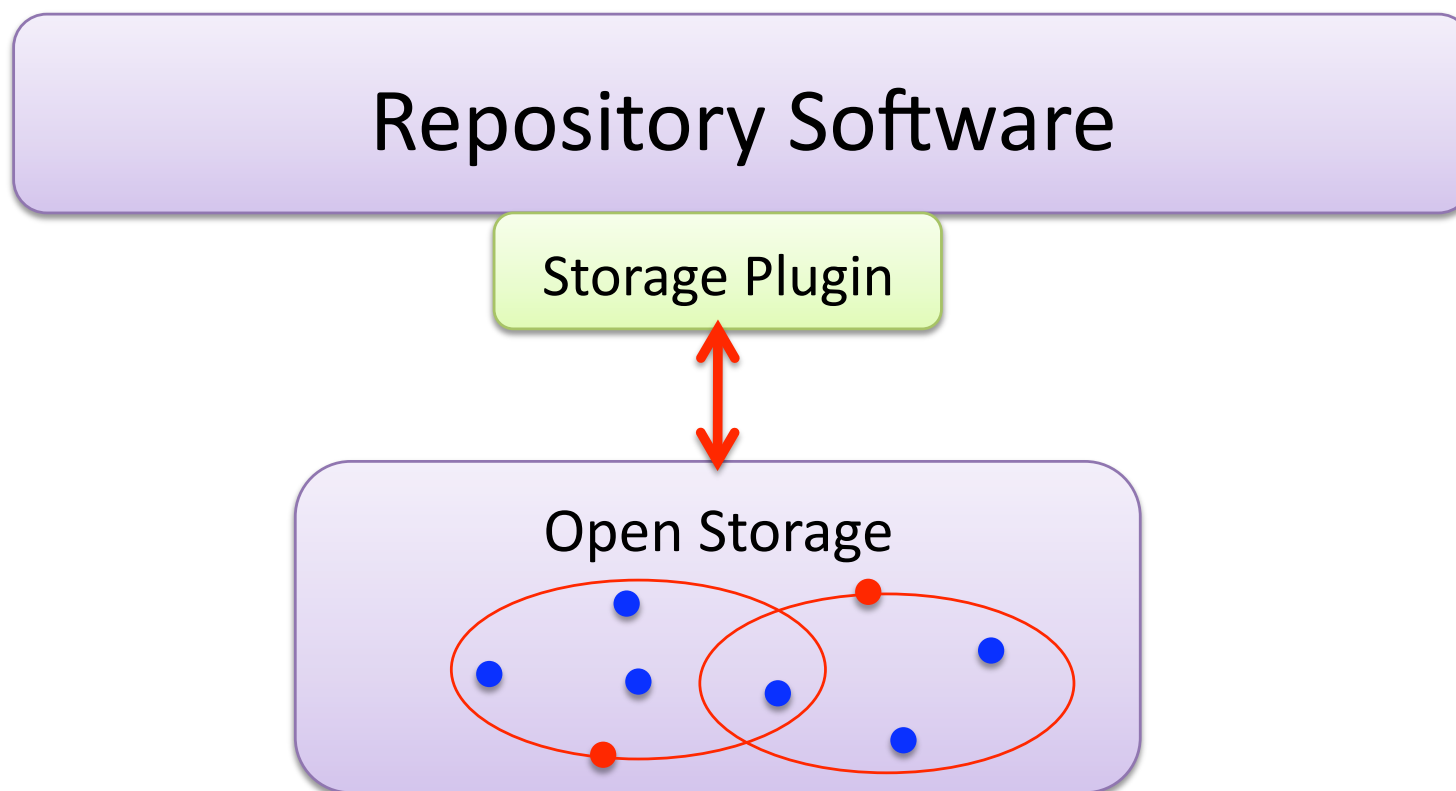
- **Reliable**
 - Self Checking and Self Healing file System
- **Resilient**
 - Must have the capability to be robust in the case of part failure
- **Simple & Expandable**
 - Must be made of parts which are easy to expand / upgrade way into the future.
- **Open**
 - Any software developed to enable all of the above must be open, same with any hardware specifications.

Open Storage (2)

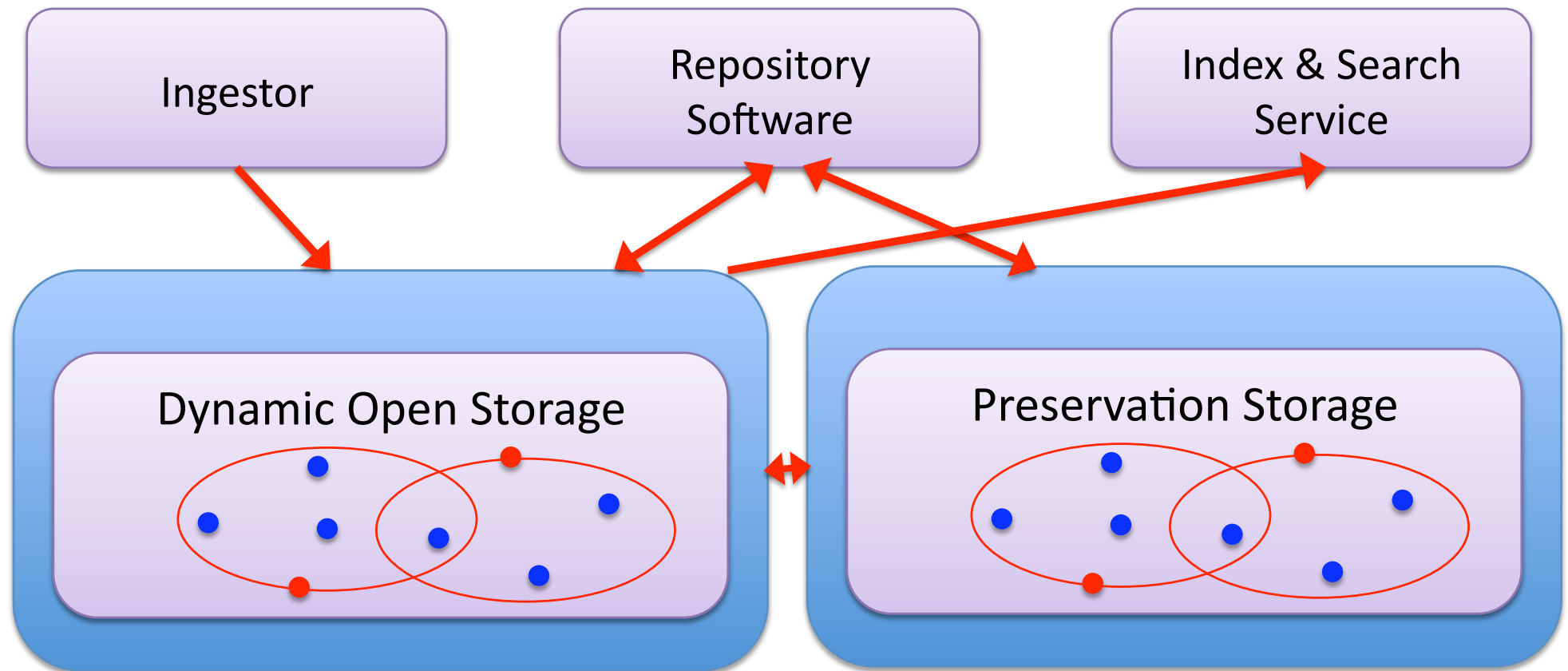
- Objects are addressable via HTTP
 - <http://url/uuid>
- Each object should be heavily tied to its UUID; you should be able to find the UUID from the object.



Repositories and Open Storage



Services over Storage



2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries

Summary

- Objects, including REMs are stored in Open Storage.
- Web Services like OAI-PMH will still be used.
- The repository software can be removed, replaced and more easily upgraded.
- Enables preservation as you can let a 3rd party have some say in the management of your objects.

Open Storage & The Cloud

- Open Storage provides a controlled web based environment in which objects can be stored.
- Still in the Cloud
- Reliable, Resilient, Expandable.

Unlocking the Repository

Repository as a service rather than software
running on a local 'box'

Facilitates a greater number of services to
operate directly with your resources

LOCKSS

Format Classification

Format Migration

2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries



“The coolest thing to do with your data will be thought of by someone else!”



2nd European Workshop on the Use of Digital Object
Repository Systems in Digital Libraries

One more thing...

- Smart storage.
- Utilising the masses of processor power in the storage cloud to benefit your repository.
- Storelets on Honeycomb

Thanks

JISC



<http://www.preserv.org.uk>

David Tarrant
University of Southampton
dct05r@ecs.soton.ac.uk

Questions

- But how do I control access to my objects?
 - Use n storage plugins and n indexers and then perform federated search.
 - Authentication
 - Use Open Auth
 - Local Login Servers
 - Eduroam
- Complete control costs \$\$\$

Questions

- Q: What happens if you inherit a single hard disk with a series of files on it but don't know any of the file identifiers. You have ORE maps for the resources. How do you identify the files which go with the URIs?
- A: When droid or an identification tool classifies the file It outputs XML which contains a binary stream of the start n bytes of the file it relates to, or a hash of the file. Smart storage!