



EPrints

10 Years of Digital Preservation



What is EPrints For?

- EPrints offers a safe, open and useful place to store, share and manage material in the pursuit of research and educational agendas.

administrative reporting, collaboration, data sharing, digital profile enhancement , e-learning, e-publishing, e-research, marketing, **open access**, preservation, publicity, research assessment, research management, scholarly collections



Why?

Providing a designated, managed and curated environment with institutional backing, which is what an IR should be, is better than not.



Why

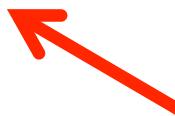
- ▶ In the late 1990s researchers started putting stuff online.
 - ▶ No management
 - ▶ No long term strategy
 - ▶ No editorial rights
 - ▶ No clue of impact
- ▶ From this was born OAI and the first EPrints repository.



Why

- ▶ In the late 1990s researchers started putting stuff online.
 - ▶ No management
 - ▶ No long term strategy
 - ▶ No editorial rights
 - ▶ No clue of impact

These are all institutional problems





Research Problems

- Discoverability
- Manageability
- Value Tracking



EPrints 1.0

- ▶ A centrally controlled IR
- ▶ Interfaces for:
 - ▶ Depositors
 - ▶ Editors
 - ▶ Managers
 - ▶ Depositors
- ▶ Indexable by search engines, download tracking
- ▶ Centrally controllable resources, easier to **back up**



- ▶ What is in the repository?

- ▶ <http://roar.eprints.org>
- ▶ PRONOM-DROID

- ▶ Applying OAIS, PREMIS.

The **technical registry**
PRONOM



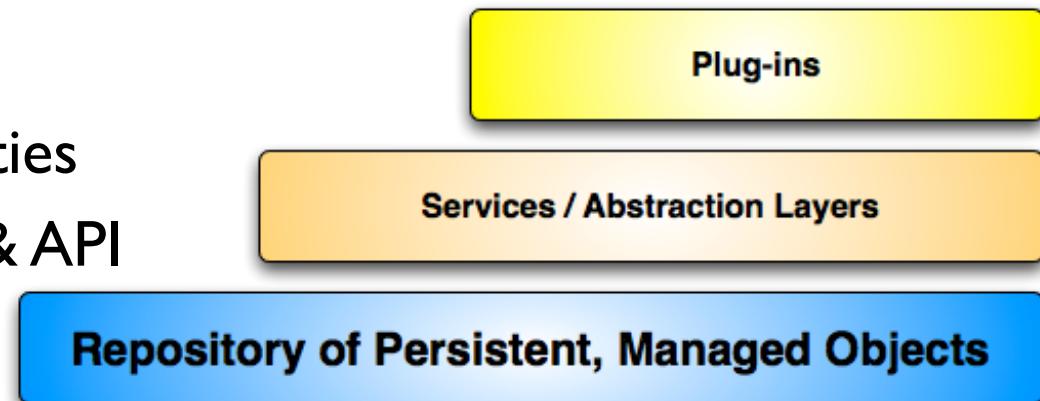
Preserv2 - 2007

- Extended preservation services
 - Preserv1 profile in the repository.
 - Storage Services for greater flexabiioty in bitstream Preservation
- Repository Interoperability (OAI-ORE)
 - First real world use of ORE.



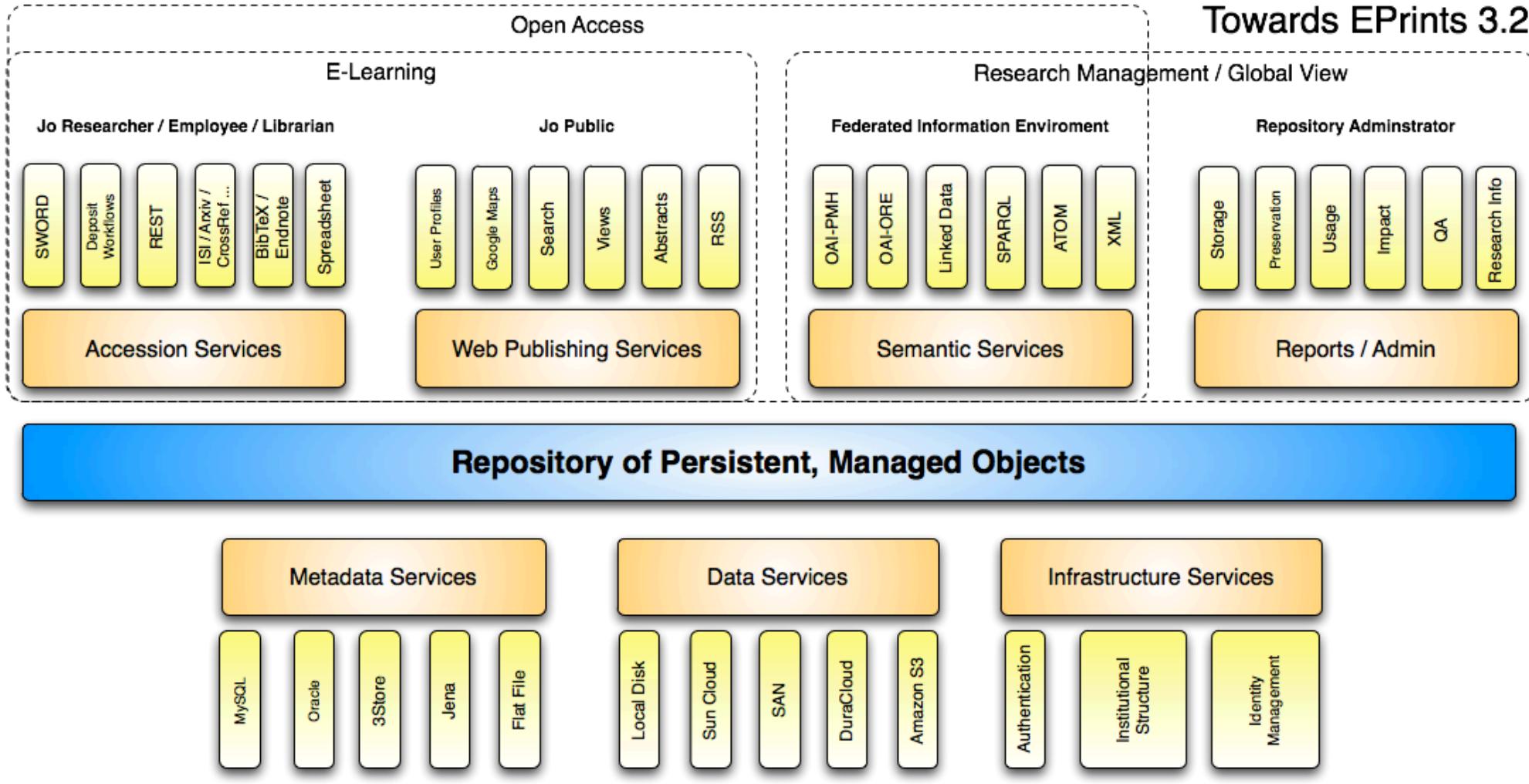


- ◀ Plug-ins / Modules
- ◀ Everything builds on the core layer
- ◀ Major part of v3.2 is strengthening the core and adding more abstraction layers
 - ◀ Improved data model
 - ◀ Enhanced data facilities
 - ◀ Enhanced metadata facilities
 - ◀ Improved programming & API





EPrints 3.2 Structure





EPrints Plug-In Biodiversity

- ▶ Atom, RSS
- ▶ Bibliography Publishing (personal, projects)
- ▶ ISI, Web of Science, Google Scholar
- ▶ MS Office (via SWORD)
- ▶ LDAP, NIS, RADIUS (Authentication)
- ▶ OAI-PMH, OAI ORE
- ▶ Oracle, MySQL
- ▶ Amazon S3, SunCSS, NFS
- ▶ DROID, Pronom, Plato Action Plans

Demo



Demo Content

- ▶ <http://eprints.ecs.soton.ac.uk>
- ▶ Search for author “Jennings”
- ▶ re-order results by ISI citation count
- ▶ Main aim is to show ISI and Google Scholar plug-ins



Demo Content

► <http://languagebox.eprints.org>

- Show the high amount of customisation (still eprints)
- Show how collections are built
- Emphasize multi-language support as all text is phrases



Demo Content

► <http://www.edshare.soton.ac.uk/1763/>

- Click the video preview link
- Demonstration of repository handling multi-formats

► <http://ecrystals.chem.soton.ac.uk/20/>

- 3D crystallography structures
- Large binary datasets are held in dark repository backed with a honeycomb



Demo Content

- ▶ Showed Storage Controller. EPrints using Sun Cloud Storage Service. EPrints is test partner.

Local Disk Storage

Local disk storage

There are 2072 total files stored using this back-end, taking 89Mb.

| | | |
|------------|------|---|
| Documents: | 409 | <input type="button" value="Copy to"/> <input type="button" value="Delete Copies"/> |
| History: | 1663 | <input type="button" value="Copy to"/> <input type="button" value="Delete Copies"/> |

Compressed local disk storage

There are 84 total files stored using this back-end, taking 402Kb.

| | | |
|----------|----|---|
| History: | 84 | <input type="button" value="Copy to"/> <input type="button" value="Delete Copies"/> |
|----------|----|---|

Archival Storage

HoneyComb storage

There are 0 total files stored using this back-end, taking 0b.

Cloud Storage Platforms

Sun Cloud Simple Storage

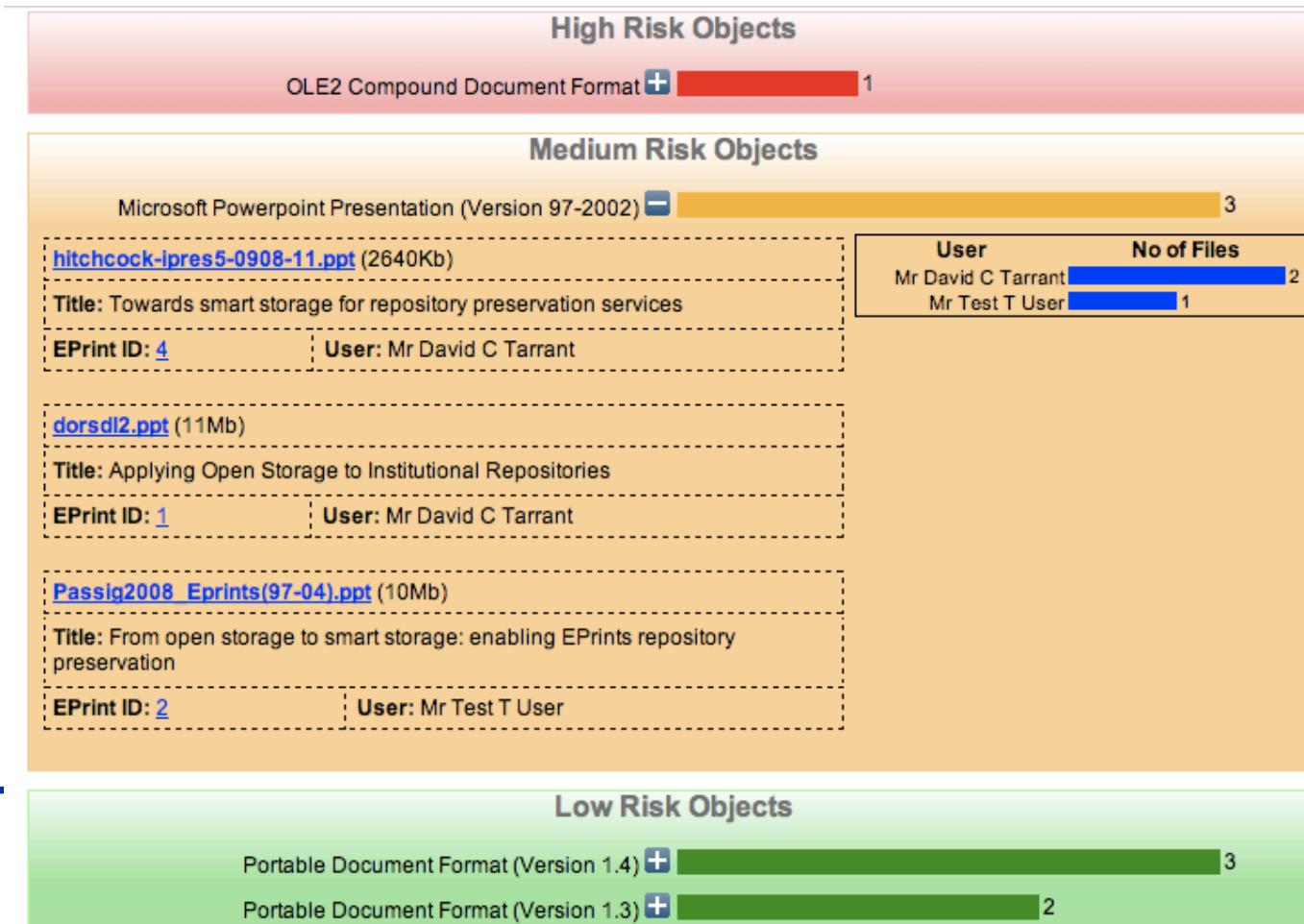
There are 12 total files stored using this back-end, taking 256Kb.

| | | |
|------------|----|---|
| Documents: | 12 | <input type="button" value="Copy to"/> <input type="button" value="Delete Copies"/> |
|------------|----|---|



Demo Content

► Preservation, file format & risk analysis plug-ins



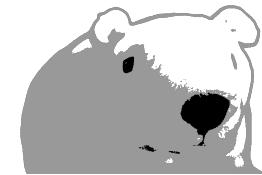


EPrints & Planets

- ◀ Classify Objects (EPrints)
- ◀ Identify Objects at Risk (EPrints)

- ◀ Load sample objects into Plato
- ◀ Analyse object properties, locate tools, test runs, analysis of results (Plato)
- ◀ Create Action Plan (Plato)

- ◀ Load Action Plan Into EPrints
- ◀ EPrints performs actions on objects (if any)
- ◀ Stores action plan for future reference of decision process.



More Info @ <http://eprints.ecs.soton.ac.uk/17962/>



<http://preservation.eprints.org>





- ▲ Aim: To create a number of exemplar preservation repositories from which others can learn
- ▲ Small number of very diverse repositories



Training



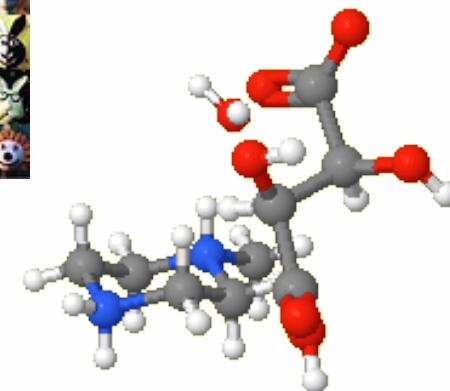
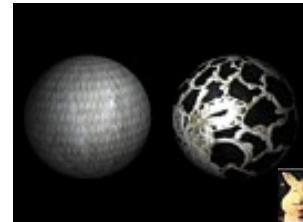
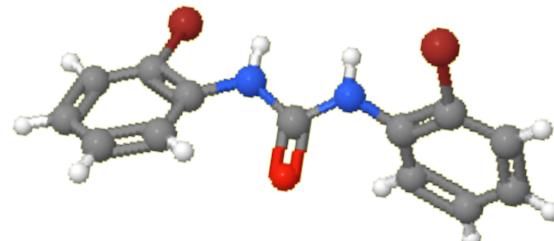
Deployment



Development



Biodiversity





Takeaway Messages

- ▶ Modular, modular, modular
- ▶ Flexible, Extensible
- ▶ Secure, Managed, Reliable
- ▶ Supported

- ▶ Embrace change
- ▶ Support your community



davetaz@ecs.soton.ac.uk

THANK-YOU