

DIRECTIONS FOR EPRINTS

Dr Leslie Carr
University of
Southampton,
Web Science
Institute
& EPrints
Services

<http://www.flickr.com/photos/coderkind/408839789/>

Decades of Achievement

Internet (packets)

Web (documents)

Repository (preservation)



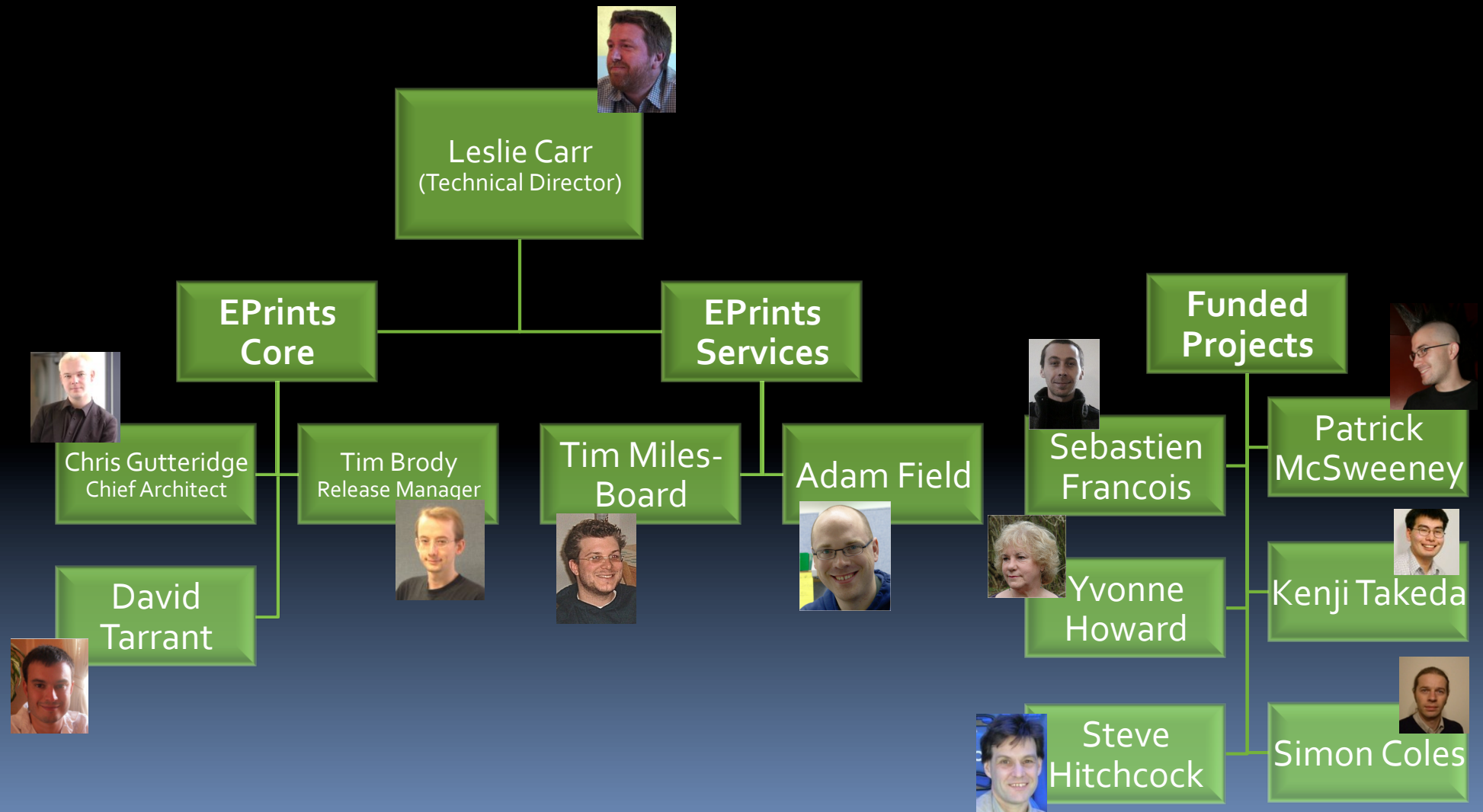
What have we achieved with this multilayered platform in the last decade?
What scope have we for achieving new things?



EPrints History

- Conception
 - ▣ Oct 1999
- Birth
 - ▣ June 2000
- Version 2
 - ▣ Feb 2002
- Version 3
 - ▣ Jan 2007

EPrints Org Chart



EPrints Services

- For-pay commercial hosting, bespoke development, training, support
 - two fulltime repository developers/supporters
- Provides resources to sustain EPrints O/S development
- Provides customer contact
- Contributes code back to O/S
- Funds a PhD studentship
 - Bibliographic & social network analysis

What is the Point?

- Universities and researchers are knowledge producers and knowledge consumers
- The Web has radically altered the potential for knowledge dissemination in society in ways that we haven't yet fathomed
- We want to understand and facilitate that change
 - Research *and* development

Openness Agendas

- Open Access to Research Outputs
 - Supported by all Research Councils
 - Heavily promoted by JISC
- Open Educational Resources
 - Move from “Learning Objects” to “Learning Resources” via Open Access practice and methodology
 - Move from closed VLEs to sharing resources
 - All underpinned by active preservation and curation policies and workflows
- Open Research Data
 - Move from closed VRE / EScience environment
 - To open (able) data



Open Access to Research Outputs

- “A new public good”

Proteomic analysis of cell models of ovarian cancer tumour suppression
Sinclair, J.R. (2010) Proteomic analysis of cell models of ovarian cancer tumour suppression. Doctoral thesis, UCL (University College London)

[PDF](#) - Requires a PDF viewer such as [GSview](#), [Xpdf](#) or [Adobe Acrobat Reader](#)
5Mb

[Excel Spreadsheet \(Appendix 1A: whole cell lysate TMT\)](#)
4Mb

[Excel Spreadsheet \(Appendix 1B: secreted TMT\)](#)
1434Kb

[Excel Spreadsheet \(Appendix 1C: membrane TMT\)](#)
2056Kb

Abstract
Epithelial ovarian cancer (EOC) is the most common form of gynaecological malignancy in the developed world. Identifying molecular markers of disease may provide novel approaches to screening and could enable targeted treatment and the design of novel therapies. In previous work, 141 primary ovarian tumours were analysed using metaphase comparative genomic hybridization to identify complete or partial chromosome deletions that may harbour tumour and/or metastasis suppressor genes. Chromosome 18 (Ch18) was found to have deletions in 50% of the tumours. Microcell-mediated chromosome transfer (MMCT) of normal Ch18 material into EOC cell lines resulted in tumours that displayed significant suppression of anchorage dependence, growth, invasiveness and reduced tumour growth in nude mice.

UCL LIBRARY SERVICES
UCL Eprints

Library home » Online resources » UCL Eprints

Home :: Search :: Browse :: Information :: Statistics :: Latest additions :: Contact

Welcome to UCL Eprints

Quick search
Enter search term(s) [Search](#)

Information and help
[More](#) about UCL Eprints including FAQ, copyright, deposit guide and research funders' policies.

UCL's Open Access Policy
UCL Academic Board has approved an Open Access policy, ensuring that, subject to copyright permissions, all UCL research is placed online in UCL Eprints, freely available to all. For more details, see the June 2009 [Press Release](#). A UCL Publications Board has been convened to oversee the implementation of the UCL Open Access Policy.

For further information about UCL Eprints and Open Access generally, see [About UCL Eprints](#).

Research Funders' policies on Open Access
Holders of Wellcome Trust awards should consult the UCL Library Services [Guide to Wellcome Trust open access funding](#) for advice on compliance with the Trust's grant conditions. For advice on compliance with the policies of other research funders, please visit the Library's [Research Funder Policies](#) pages.

Statistics - most downloaded

Annual	2009	::	2008
Quarterly	Last quarter	::	All
Monthly	Last month	::	All

June 2010

Eprint	Downloads
Blow, L. and Leicester, A. and Oldfield, Z. (2003) London's congestion charge. IFS Briefing Notes (BN31). Institute for Fiscal Studies, London, UK.	1219
Sinclair, J.R. (2010) Proteomic analysis of cell models of ovarian cancer tumour suppression . Doctoral thesis, UCL (University College London)	1097
Owor, M. (2010) Groundwater - surface water interactions on deeply weathered surfaces of low relief in the Upper Nile Basin of Uganda . Doctoral thesis, UCL (University College London)	453
Krashia, P. (2009) Recombinant glycine receptors: stoichiometry and kinetics . Doctoral thesis, UCL (University College London)	294
Dustmann, C. and Casanova, M. and Fertig, M. and Preston, I. and Schmidt, C.M. (2003) The	

[Del.icio.us](#) [reddit](#) [StumbleUpon](#) [Facebook](#) [Digg](#)

Open Arts

- JISC KULTUR consortium of universities and art colleges
- Capturing performance, practice, exhibition, evidence
 - In a *no publication* environment

UAL Research Online

Sample Works



Williams, Keir (2008) *Digital is dangerous debate: digital vs analogue*. [Creative Arts and Design > Film & not elsewhere classified]

University of the Arts London

Home
Browse
Latest
About
Policies
Login

Browse collection by

Year
College/Collection
Subject
Creator
Projects

Latest items

Lewis, Mark (2009)
Jeff Wall - Photographer.
In: Jeff Wall- The Complete Edition. Phaidon

I am Not What I am: The Taming of the Shrew and Twelfth Night.

Pavelka, Michael (2007) *I am Not What I am: The Taming of the Shrew and Twelfth Night* (Performance) [Creative Arts and Design > Stage Design
Creative Arts and Design > Theatre Design
Creative Arts and Design > Theatrical Wardrobe Design]

Advanced Search

Images Documents Details



12th Night: a coloured stock
12th Night: Andrew & Toby & Zany
12th Night: Andrew boxing
12th Night: Andrew's hideout
12th Night: Andrew's jig
12th Night: Antonio at pumpout
12th Night: begonia
12th Night: boxing ring
12th Night: Feste & Zany

University of the Arts London

Home
Browse
Latest
About
Policies
Login

Advanced Search

Details

Name
Mr. Hall, Edward
(Writer) Warren, Roger

Finance

University of the Arts London > Stage Design
University of the Arts London > Theatre Design
University of the Arts London > Theatrical Wardrobe Design

James W. Widdowson College of Art

James W. Widdowson College of Art is a world-leading international conservatoire who has designed over 130 productions world-wide, many of which have been new plays or new musicals. His Midsummer Night's Dream (2007) won the Manchester Evening Standard Award for Best Production and Best Design (2007). 'Noise Rider' (a two-part adaptation of Shakespeare's Henry VI trilogy) won the 2002 Barings Award for Best Touring Production and (in a later production at the Chicago Shakespeare Company) the 2004 Barings Award for Best Production and Best Design.

Capturing artistic experimentation & performance vs capturing scientific experimentation

Open Educational Resources

- JISC EdSpace
- IEEE LOM metadata vs Web 2, informally tagged
- Sufficient for discoverability and usefulness

The screenshot displays the EdShare website interface. At the top, the header includes the EdShare logo, navigation links (Home, Browse, About, Help), a search bar, and user links (Profile, Share manager, Bookmarks, Saved searches, Logout). The main content area features a poster titled "What are the actual effects of Enterprise 2.0?" by Mark Schuler. The poster includes sections on "Origins of social software", "Enterprise 2.0 Typology", "Research methods", and "Theoretic framework". Below the poster, there is a list of related items with download links: "Mark Schuler poster.pptx", "Saroah Khan poster.ppt", "Michael Yip poster.pdf", and "Mike Santer.pptx". To the right of the poster, there is a sidebar with sections: "Description" (explaining the poster's origin), "Advice for reuse" (facilitating discussions), "Share details" (listing added by, on, creators, tags, viewing permissions, link, and downloads), and a "Toolbox" with options like "Edit item", "Bookmark this item", "Download as zip", and "Use as template".

Open Scientific Data

A repository of chemical data. Scientific, not bibliographic, metadata.

JISC EBank project (2003)

Run by the UK Crystallographic Service and latterly an international consortium.

eCrystals

Home | About | Browse by Year | Browse by People | Login | Create Account

University of Southampton

(6-Thien-2-ylpyrid-3-yl)methanol

Sample Originator: Donated Sample^a.

Data Collection: Susanne L. Huth^b and Michael B. Hursthouse^b

Structure Determination: Susanne L. Huth^b.

Sample kindly provided by Maybridge^a
University of Southampton^b

C₁₀H₉NOS

InChI=1/C10H9NOS/c12-7-8-3-4-9(11-6-8)10-2-1-5-13-10/h1-6,12H,7H2

Identification Number: 10.3737/ecrystals.chem.soton.ac.uk/373

Date Created: 09 July 2007

Deposited On: 31 Jan 2008 13:08

Deposited By: Ms L.S. Huth

Depositor Comments

More information about this compound can be found at:
<http://www.maybridge.com/>

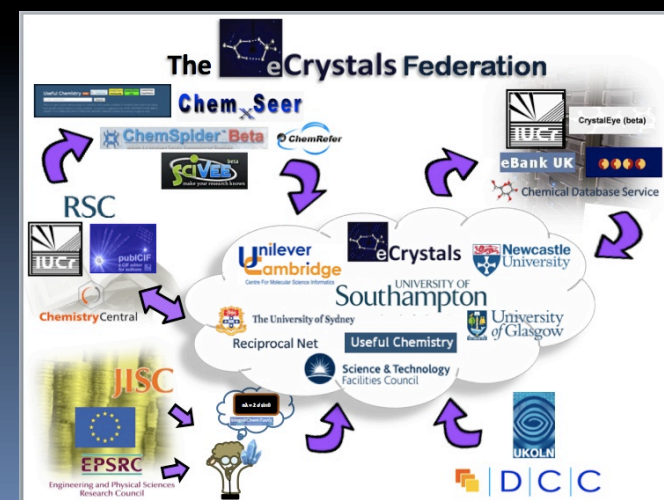
Data collection parameters

Chemical formula	C10 H9 N O S
Crystallisation Solvent	
Crystal morphology	Needle
Crystal system	Orthorhombic

Available Files

Final Result		
2007may0014.cif		11k
2007may0014.cml		3k
2007may0014.fcf		99k
Validation		
2007may0014_checkcif.htm		8k
Refinement		
2007may0014.res		4k
2007may0014_xl.lst		30k

Chemical structure image: (6-Thien-2-ylpyrid-3-yl)methanol



(Other Open Research)

- JISC myExperiment
- Capturing
 - Workflows
 - Data
- Comes from EScience/
Cyberinfrastructure
- But uses Web2 as
design principle

The screenshot displays the myExperiment website interface. At the top, there's a navigation bar with links for 'About', 'Mailing List', 'Publications', 'Log in', 'Register', 'Give us feedback', and 'myExperiment'. Below this is a secondary navigation bar with 'Home', 'Users', 'Groups', 'Workflows', 'Files', and 'Packs'. A search bar is located in the center of this bar.

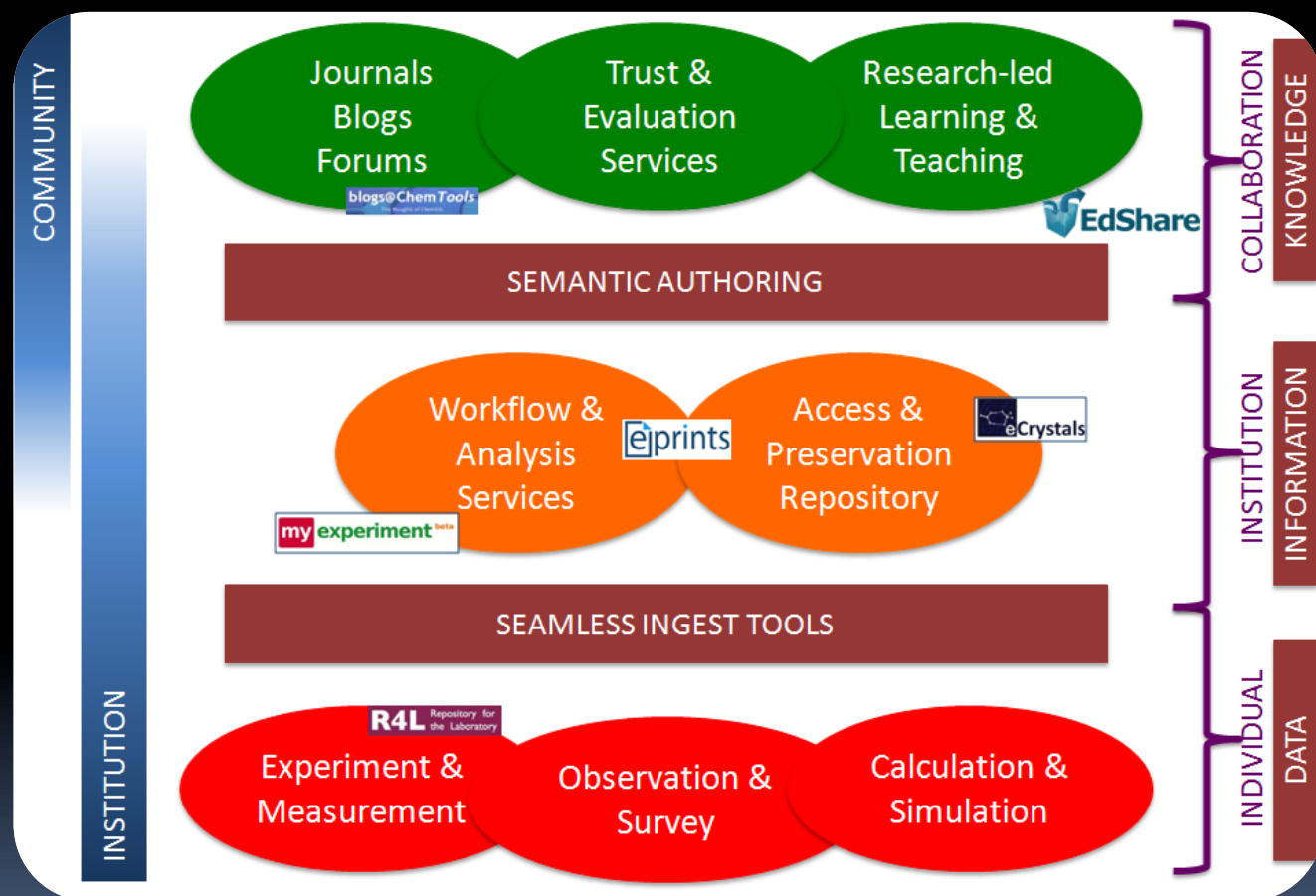
The main content area features a 'Workflow Entry: Lymphoma type prediction based on microarray data'. It includes metadata such as 'Created at: 23/04/09 @ 17:10:37', 'Last updated: 11/05/10 @ 19:04:33', and various statistics like 'License: Credits (3)', 'Attributions (0)', 'Tags (11)', 'Featured in Packs (0)', 'Ratings (0)', 'Attributed By (0)', 'Favourited By (0)', 'Citations (2)', 'Version History', 'Reviews (0)', and 'Comments (0)'.

The workflow is shown in a 'Preview' section, displaying a complex flowchart with nodes and arrows. A 'Download Scalable Diagram (SVG)' button is at the bottom of the diagram.

On the right side, there's a sidebar with several sections: 'New/Upload' with a 'Workflow' button and a 'GO' button; 'Log in / Register' with fields for 'Username or Email:', 'Password:', and 'Remember me: ☐', and a 'Log in' button; 'Need an account? Click here to register' and 'Forgot Password?'; 'Popular Tags' with a list of 25 tags including 'benchmarks', 'bioinformatics', 'BLAST', 'd-grid', 'data integration', 'ebi', 'example', 'graph', 'gwas', 'gworkflow', 'kegg', 'localworker', 'microarray', 'mygrid', 'ondex', 'ovtk', 'oxl', 'pathway', 'pathways', 'protein', 'pubmed', 'rdf', 'taverna', and 'text mining'; and 'Original Uploader tags' with a list of tags including 'cabig', 'cagrid', 'cancer', 'research', 'classification', 'gene', 'expression', 'globus', 'grid', 'service', 'lymphoma', 'lymphoma', 'workflow', 'microarray', 'support vector machine', and 'Log in to add Tags'.

Understanding Data Management

- JISC Institutional Data Management Blueprint (IDMB)
- Campus-wide study to produce a 10-year plan



**The challenges of human scale
institutional repositories versus the
challenges of industrial-scale
processing of humongous collections.**

Lawnmowers vs Combine Harvesters?

**HOW DO YOU MANAGE AN ENTIRE
NATION'S GRASS CLIPPINGS?**

Flashback to 2007: a slide originally about a grassroots approach
to the preservation of scholarly material

EPrints Activity

Research Business

Research
Management

Research
Processes

Research
Impact

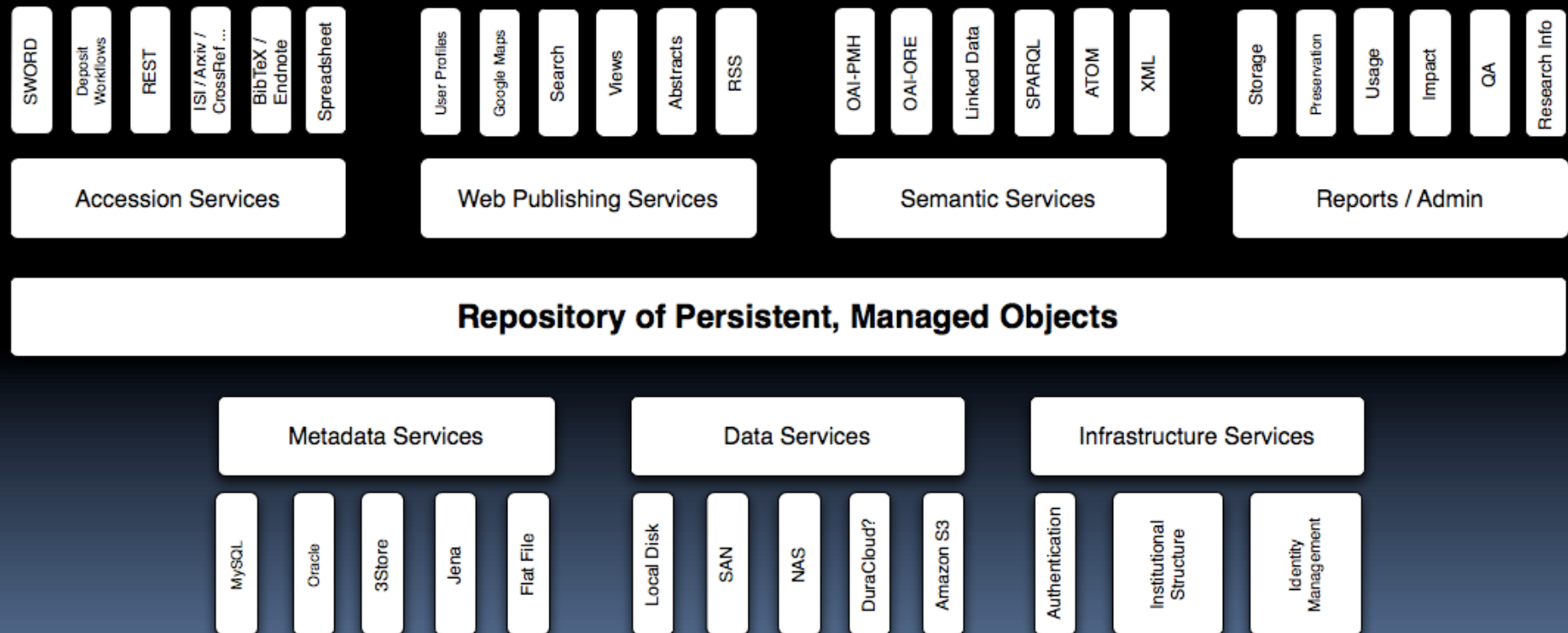
Research Content

Research
Publications

Educational
Resources

Scientific Data

EPrints Software Architecture

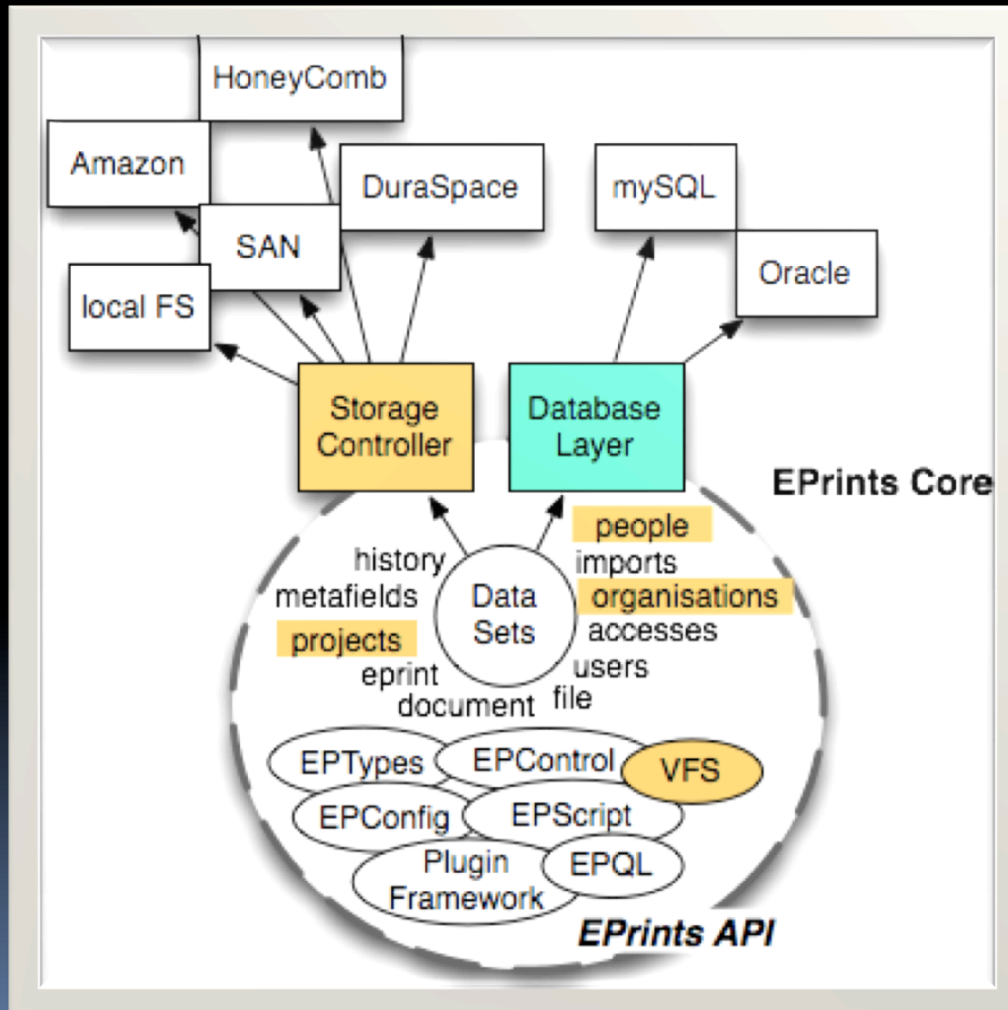


EPrints Core

Introduced in 3.2

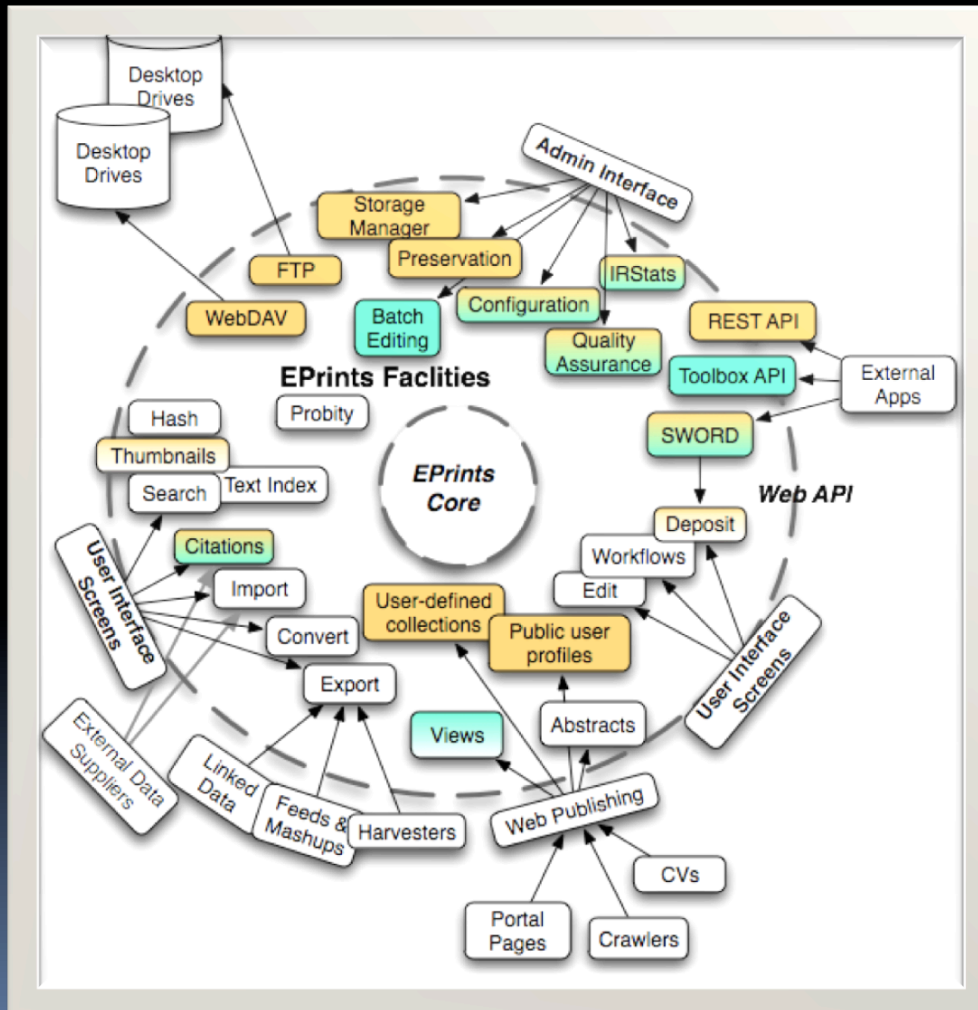
Introduced in 3.1

Introduced in 3.0



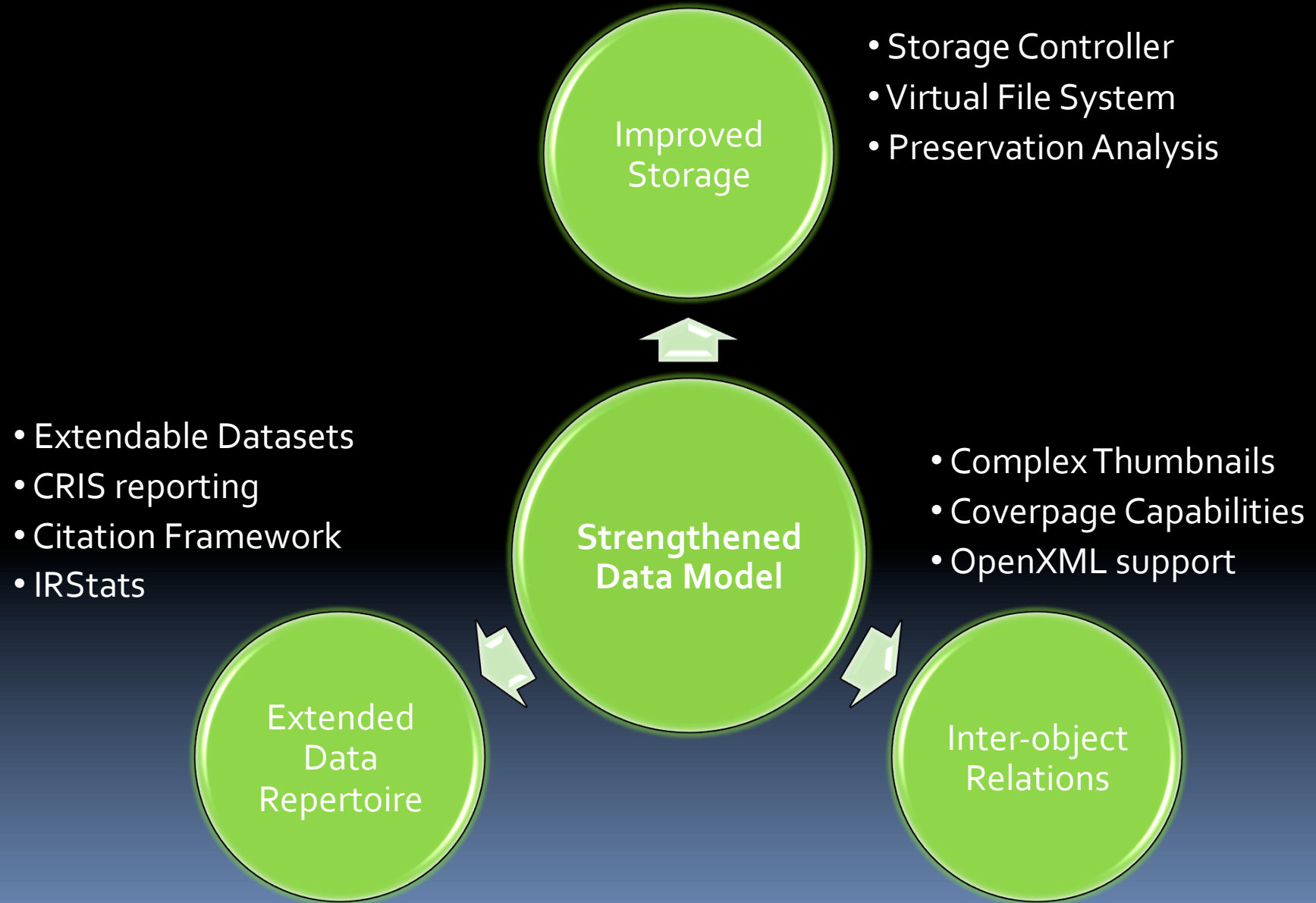
- Datasets
- Data storage
- Utility methods
 - Data types
 - Configuration control
 - Query language
 - Control language *etc*

EPrints Mantle



- External interactions
 - web pages
 - web services
 - other protocols

EPrints 3.2 Improvements



EPrints 3.2 Improvements

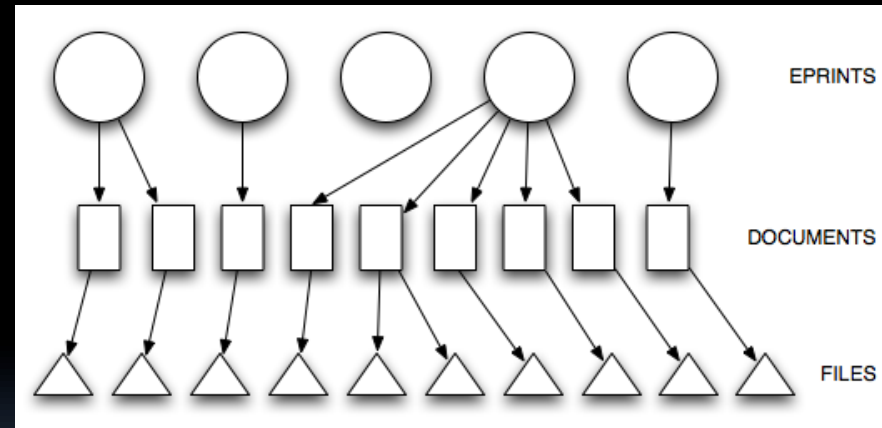
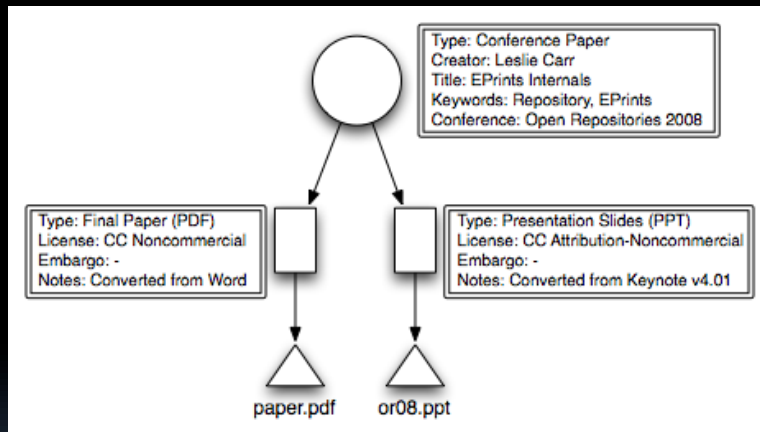
Database	API	Documents	Deposit Interface	Search & Indexing	Unicode
REST	SWORD ₂	Linked Data Support	Collections Support	Semantic Web / Linked Data (RDF)	Storage Layer
Speed	Import	EPC & EPrints Script	OAI	Unit Tests	Metadata Types
Administration Interface	Editorial Interface	User Defined Datasets	Command Line Tools	Export	Abstract Page
Phrases		Views	Key Bug Fixes		

See EPrints Wiki for details – http://wiki.eprints.org/w/New_Features_in_EPrints_3.2

Improved
Storage

Files are now data objects

- Previously, eprints & documents were data objects. Files were just items in a file system.



- Now files are data objects too, and file objects are abstractions of data streams.



Improved
Storage

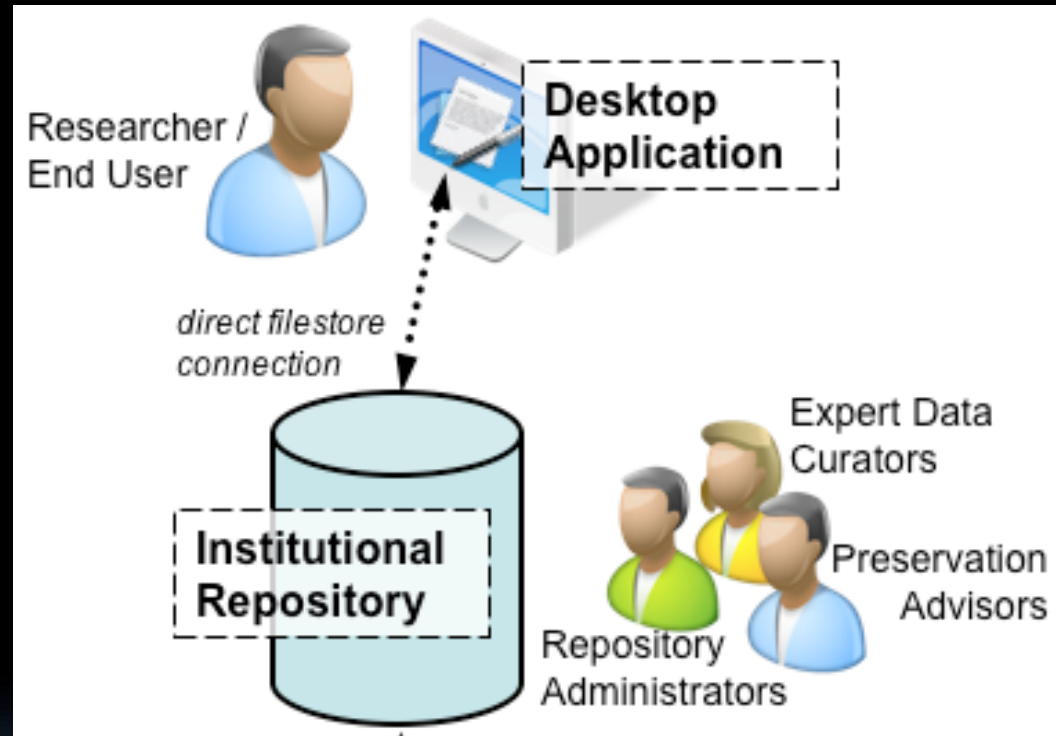
Storage Controller

- Files can be stored on multiple storage services
 - ▣ Local disk, SAN, NAS, Honeycomb, Cloud
- Various policies can be enacted through an EPControl XML files.
- Services can be monitored and files swapped between services.

Improved
Storage

Virtual File System

- A VFS module invents a view of the repository's files and structures as a hierarchical file system.

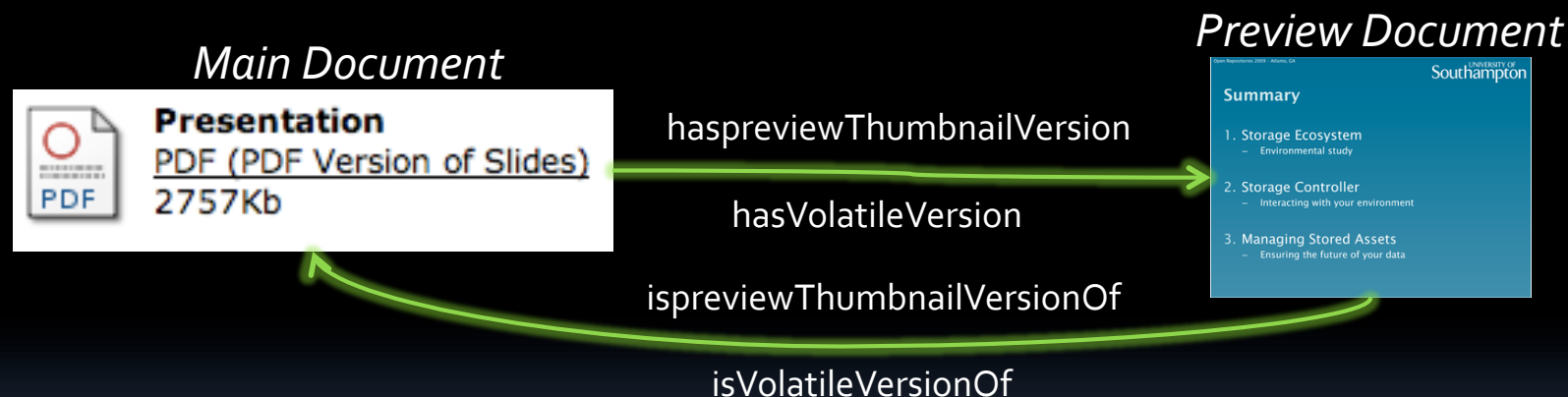


- *WebDAV and FTP servers are implemented to allow client desktops to mount the repository as a pseudo-filesystem.*



Complex Thumbnails

- Thumbnails were hidden files
- Now they are documents



- Now thumbnails can be different formats
 - png, jpg, flv...
- Thumbnails can be uploaded manually

Thumbnail Addressing

- Main document
 - <http://eprints.com/123/4>
- Thumbnail
 - <http://eprints.com/123/4.haspreviewThumbnailVersion>
- Chaining Relationships
 - <http://eprints.com/123/4.haspreviewThumbnailVersion.ispreviewThumbnailVersionOf>

Object IDs, REST & Linked Data

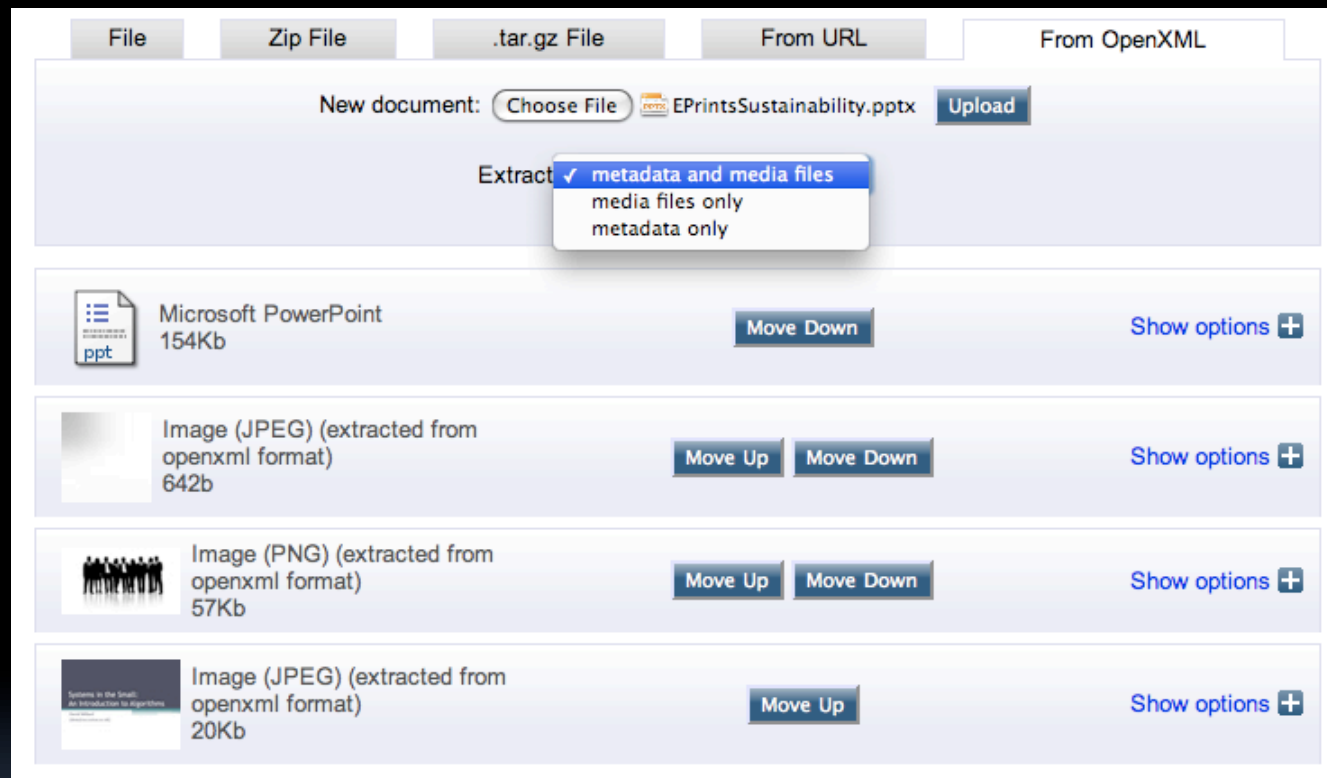
- All data objects have their own data set and id
 - <http://eprints.com/id/eprint/123>
 - <http://eprints.com/id/user/123>
 - <http://eprints.com/id/document/123>
- Type these into a browser URL and EPrints will redirect to the appropriate display object
 - e.g. the eprint that contains document id 123
- Or in other contexts to RDF or other

Object IDs & REST

- EPrints objects are exposed as data for reading and writing through a Web interface
- <http://devel.eprints.org/cgi/rest/eprint/4.xml>
 - The XML output of an eprint
- <http://devel.eprints.org/cgi/rest/eprint/4/creators/2/name/family.txt>
 - text of the surname of the second author of eprint 4

Interobject Relations

Microsoft Office Support



- A new document upload tab allows PPTX/DOCX files to be burst open
 - Media files stored as related documents
 - Metadata added to the eprint
- Also happens with SWORD plugin

Demos

Coming in 3.3+

- Final CERIF extensions
- EPrints Package Manager / App Store
 - ▣ Launch at Repository Fringe, Sep 2010
- Scheduler / Calendar interface to event planning
- Multistage Editorial Control
- Manual raising of Quality Issues

Conclusions

- EPrints as a competent and mature part of the researcher's information environment
- Integrating with their desktop knowledge creation activities (making and editing documents)
- Syndicating rich summaries to Web portals, Web 2 and social networking environments
- Participating in linked data / semantic web

Discussion & SWOT Analysis

Strengths

Weaknesses

Opportunities

Threats