

Repositories for research – contributing to the Knowledge Cycle

Pauline Simpson
National Oceanography Centre, Southampton

JIBS Conference
15 Sep 2005
British Geological Survey, Nottingham



**National Oceanography
Centre, Southampton**
UNIVERSITY OF SOUTHAMPTON AND
NATURAL ENVIRONMENT RESEARCH COUNCIL



National Oceanography Centre, Southampton



NOC is one of the world's leading centres for research and education in marine and earth sciences, for the development of marine technology and for the provision of large scale infrastructure and support for the marine research community

Joint Venture between Natural Environment Research Council and the

University of Southampton

Research-led multidisciplinary university:

20,000 students

5000 staff (3000 researchers)



Outline

- Open Access Context
- The Southampton Route
- Building on Institutional Repositories
 - Contributing to the Knowledge Cycle

Open Access

- Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions.
- OA should be immediate, rather than delayed, and OA should apply to the full-text, not just to abstracts or summaries.
- OA removes *price barriers* (subscriptions, licensing fees, pay-per-view fees) and *permission barriers* (most copyright and licensing restrictions).

See JISC briefing paper on Open Access April 2005

http://www.jisc.ac.uk/index.cfm?name=pub_openaccess

Open Access to Research

- 27 Jun 1994 Stevan Harnad's 'Subversive Proposal' leading to the open access vision for scholarly material
 - ("Faustian Bargain" with publishers – a price tag barrier to research)
 - **Harnad, S.** (1995) A Subversive Proposal.
In: Ann Okerson & James O'Donnell (Eds.) Scholarly Journals at the Crossroads: a Subversive Proposal for Electronic Publishing. Washington, DC., Association of Research Libraries, June 1995.
<http://www.ecs.soton.ac.uk/~harnad/subvert.html>
<http://www.arl.org/scomm/subversive/toc.html>
 - In an ideal world of scholarly communication – all research should be freely available -



But journals still becoming more and more expensive (serials crisis)

- journals are the primary research publication channel
- journal publishing is dominated by commercial ventures
 - Researchers write papers for journals (free or page charges!)
 - Researchers transfer copyright to publishers (free)
 - Researchers on Editorial Board (free)
 - Researchers review papers (free)

BUT

- Libraries pay huge subscriptions to publishers to access the paper (and electronic) and universities pay more than once: subscription, photocopying license and for study packs
- Or possibly they cannot afford the subscription

1986-2000
Journal price
inflation +291%
Retail price index
+ 70%

Two Complementary Solutions

- **Open Access Repositories** articles, conference papers, reports, theses, learning objects, multimedia etc. deposited in open electronic archives/repositories which conform to

Open Archives Initiative (OAI) standards

- **Open-Access Journals:** Journals will not charge subscriptions or fees for online access. Instead, they should look to other sources to fund peer-review and publication (e.g., publication charges)
 - *No payment by author = open access or subscription*
 - *Publishing model – author pays = OA*

Repository Development

ArXiv (from 1991 at Los Alamos now at Cornell) for high energy physics community (incl Atmospheric and Oceanic Physics, Math, Computing Science and Nonlinear Science).

Despite success of ArXiv and others - RePEc (Economics), Cogprints (Cognitive Psychology), Mathematics, etc – varying success by other subject communities (Chemistry Preprints Server now finished)

2000 onwards complementary implementation of Institutional repositories fuelled by project funding eg Mellon Foundation, Howard Hughes, Open Society Institute, JISC UK FAIR Programme and powered by the Information Community

Institution Archives Registry

[Home](#) [Browse](#) [Analysis](#) [Register an Archive](#) [About](#) [Login](#) [Register Institutional Self-Archiving Policy](#)

Any Geographic Country Any System Software Any Content Type

Page 1 of 47 (466 archives) [1 2 3 4 5 6 7 8 9 10] [View All]

Name	Summary Graph	Thumbnail
<p>1) 11th Joint Symposium on Neural Computation GNU EPrints (OAI) United States e-Journal/Publication info:other.archives.eprints.org/import Total OAI Records: 30 100% freely accessible fulltext (estimate)</p>		
<p>2) Acadia Divinity College - Digital Collections DSpace (OAI) Canada Other info:other.archives.eprints.org/import Archive of lectures Total OAI Records: 232 100% freely accessible fulltext (estimate)</p>		
<p>3) Advanced Knowledge Technologies (AKT) EPrints Archive GNU EPrints (OAI) United Kingdom Research Cross-Institution Total OAI Records: 195 100% freely accessible fulltext (estimate)</p>		
<p>4) AgEcon Search: Research in Agricultural and Applied Economics Other softwares (various) United States Research Institutional or Departmental 100% freely accessible fulltext (estimate)</p>	<p>OAI Interface Unknown</p>	
<p>5) AgentLink.org EPrints AgentLink Publications Clearinghouse GNU EPrints (OAI) United Kingdom Research Cross-Institution Total OAI Records: 1197 100% freely accessible fulltext (estimate)</p>		
<p>6) ALADIN :Accès Libre aux Archives du Dépôt Institutionnel Numérique de la MSH-Alpes</p>		

Country

	Archives	In Celestial	Records	Mean	Median
• United States	125	96	584682	6090	239
• United Kingdom	54	43	89636	2085	192
• Germany	40	31	53580	1728	345
• Brazil	29	20	72738	3637	119
• Canada	27	26	25830	993	159
• France	23	19	49229	2591	181
• Italy	19	15	8234	549	100
• Australia	18	15	57750	3850	305
• Netherlands	16	12	306510	25543	3689
• Sweden	12	8	23948	2994	3284
• India	11	6	3383	564	366
• Spain	8	5	52493	10499	516
• Japan	6	6	79382	13230	806
• Belgium	6	4	2051	513	858
• Commercial	5	3	231	77	61
• Portugal	5	3	2211	737	338
• Hungary	4	4	1662	416	827
• South Africa	4	2	247	124	63
• China	4	2	13923	6962	5927
• Denmark	4	3	3996	1332	896
• Switzerland	4	3	76395	25465	5682
• Finland	4	4	12769	3192	6033
• Colombia	3	1	2335	2335	2335
• Mexico	3	3	125174	41725	4272
• Austria	3	3	2003	668	667
• Russian Federation	2	2	683	342	242
• Ireland	2	2	275	138	120
• Norway	2	1	204	204	204
• Chile	2	1	407	407	407
• Greece	2	2	46498	23249	11928
• Educational	1	?	?	0	0
• Pakistan	1	1	204	204	204
• Argentina	1	1	39	39	39
• Peru	1	?	?	0	0
• Taiwan	1	1	6826	6826	6826
• Slovenia	1	1	122	122	122
• Namibia	1	?	?	0	0
• Singapore	1	1	500	500	500
• Israel	1	?	?	0	0
• Turkey	1	?	?	0	0
• Other	9	8	1478459	184807	186646

<http://www.openoar.org/>

OpenDOAR - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Refresh Mail Print

Address <http://www.openoar.org/> Go Links

OpenDOAR

DIRECTORY OF OPEN ACCESS REPOSITORIES

[Home](#) | [About](#) | [FAQ](#) | [Contact us](#)

The Directory of Open Access Repositories - OpenDOAR

University of Nottingham, UK and University of Lund, Sweden are developing a new service for Open Access to research information.

A new service is starting development to support the rapidly emerging movement towards Open Access to research information. The new service, called OpenDOAR, will categorise and list the wide variety of Open Access research archives that have grown up around the world. Such repositories have mushroomed over the last 2 years in response to calls by scholars and researchers worldwide to provide open access to research information.

OpenDOAR will provide a comprehensive and authoritative list of institutional and subject-based repositories, as well as archives set up by funding agencies - like the National Institutes for Health in the USA or the Wellcome Trust in the UK and Europe. Users of the service will be able to analyse repositories by location, type, the material they hold and other measures. This will be of use both to users wishing to find original research papers and for third-party "service providers", like search engines or alert services, which need easy to use tools for developing tailored search services to suit specific user communities.

The project is a joint collaboration between the University of Nottingham in the UK and the University of Lund in Sweden. Both institutions are active in supporting Open Access development. Lund operates the [Directory of Open Access Journals](#) (DOAJ), which is known throughout the world. Nottingham leads [SHERPA](#), an institutional repository project that has helped establish Open Access archives in 20 of the leading UK research universities. Nottingham also runs the [SHERPA/ROMEO](#) database, which is used worldwide as a reference for publisher's copyright policies.

The importance and widespread support for the project can be seen in its funders, led by the [Open Society Institute](#) (OSI), along with the [Joint Information Systems Committee](#) (JISC), the [Consortium of Research Libraries](#) (CURL) and [SPARCEurope](#).

More information on the project is available on this site through the [About](#) page.

© University of Nottingham [Contact us](#)

Main Partners:

Lund University
Univ of Nottingham)

<http://www.openoar.org/about.html> Internet

start Removable Disk (F:) iamslicrome2005.ppt UMSLGSimpsonFINAL... OpenDOAR - Microso... EN 10:32 AM

Increasing numbers – Repository choices

- Subject - arXiv, Cogprints, RePEC,
- Institutional – Southampton, Glasgow, Nottingham (SHERPA), MBA UK
- National - DARE (all universities in the Netherlands), Scotland, British Library (proposal)
- National / Subject - ODINPubAfrica
- International - Internet Archive 'Universal', OAIster
- Regional - White Rose UK
- Consortia - London-LEAP (London E-prints Access Project)
- Funding Agency – NIH (PubMed), Wellcome Trust (UK PubMed), NERC
- Project - Public Knowledge Project EPrint Archive
- Conference - 11th Joint Symposium on Neural Computation, May 15 2004
- Personal – peer to peer
- Media Type - VCILT Learning Objects Repository, NTDL (Theses), e-Theses in UK
- **Publisher** – journal archives
- **Data Repositories/Archives** - NODC, BODC, DOD, JODC, BADC etc

***** Problem is that the above are not mutually exclusive

Dilemma for Researcher

- Mandates from major funding agencies now require grantees to deposit research output in a 'designated repository' or 'any'
 - Wellcome Trust (UK PubMed) - £400 million producing 3500 papers per year
 - RCUK, NIH in USA
- Where should the full text of their research be deposited
- Researcher wants to enter metadata and deposit only once
- Situation at present
 - Duplicate keying metadata into repositories of choice
 - Harvesting, but harvester is not the choice of the depositor
 - Cannot target multiple repositories with one exercise
- Does it matter where it is deposited since Google Scholar, Yahoo, Scopus , will pick it up wherever it is?

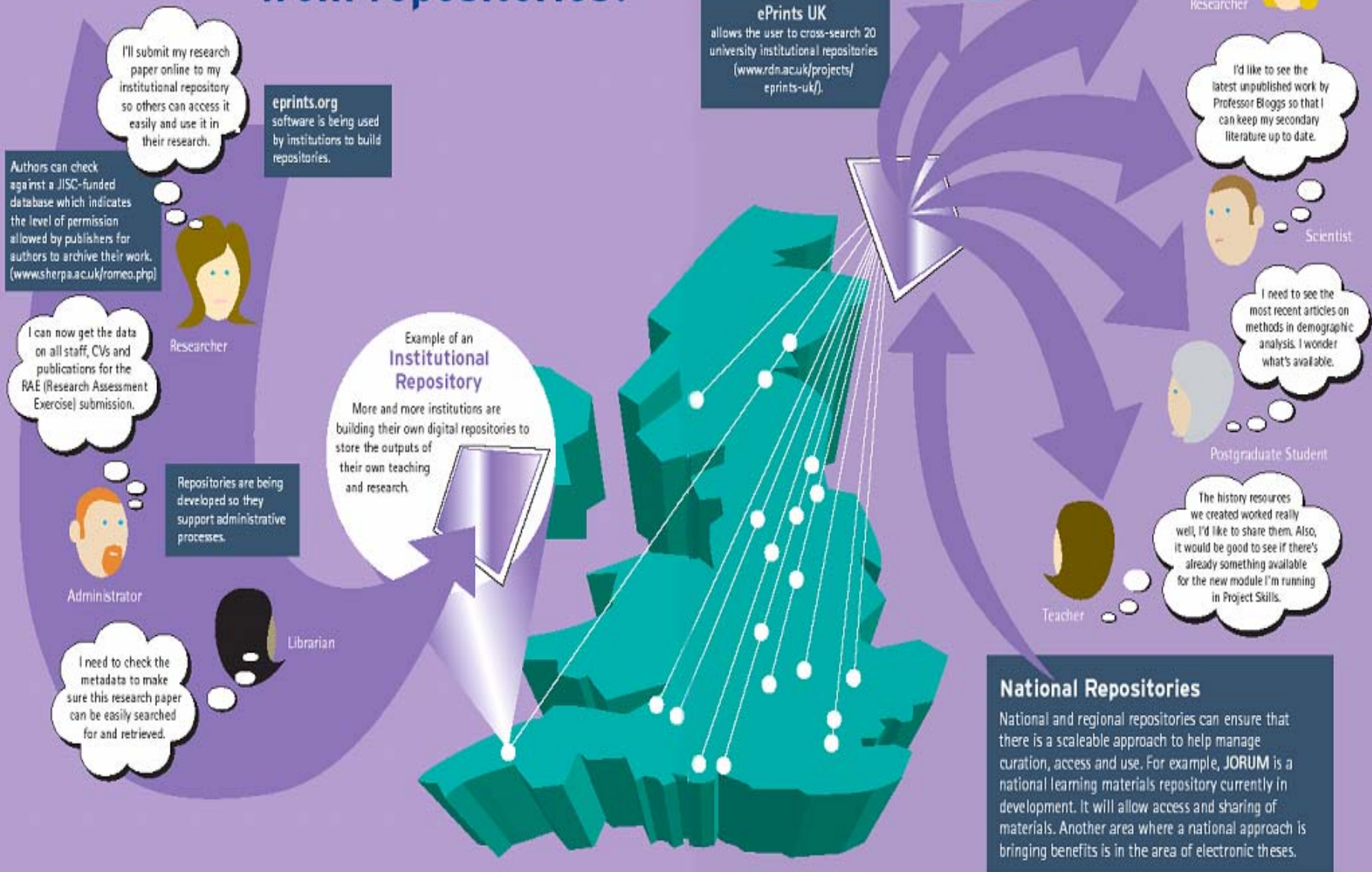
Why it should be Institutional Repositories

Institutions are logical implementers of repositories because they can take responsibility for:

- Centralising a distributed activity
- Framework and Infrastructure
- Permanence that can sustain changes
- Stewardship of Digital assets
- Preservation policy for long term access
- Provide central digital showcase for the research, teaching and scholarship of the institution

Subject or project repositories often linked to an individual or a group – can be transitory - collection at risk

Scenarios... who benefits from repositories?



Southampton's Route 2002 – 2005

built on Southampton visions

- TARDIS Project (*Targetting Academic Research for Deposit and Disclosure*) funded by JISC FAIR Programme : collaboration of The University Library, School of Electronics and Computer Science, and Information Systems Services alongside academics as one institution
 - EPrints software had been created at School of Electronics and Computer Science (ECS), Southampton to enable Harnad's self archive vision.
 - ECS already used the software for a publications database – now a sustainable repository
 - National Oceanography Centre was an early e-Print adopter

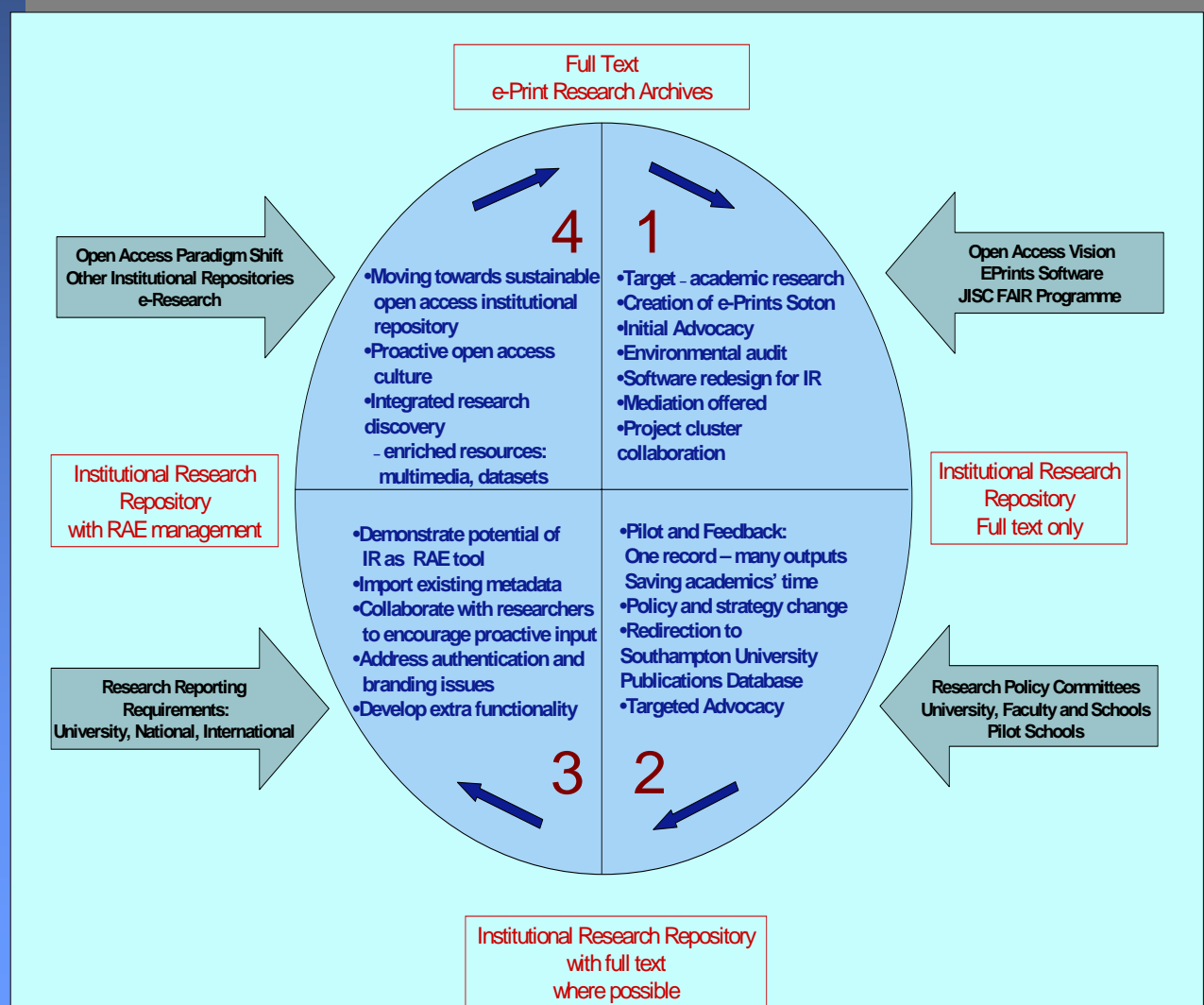
TARDIS (e-Prints Soton) Progression

- Original intent to provide secure storage and access to the full text of Southampton research output (including post refereed pre published versions of papers deposited by researchers)
 - Technical, Cultural and Organization barriers
 - Offer Mediated deposit options
- Feedback: from our advocacy, pilot and full service was that e-Prints Soton would provide maximum benefit if the service also assisted researchers with time consuming research metrics
- Evolved to 'hybrid' publications database for all research output with full text where available

Evolution from Project to e-Prints Soton : Achieving a slower but more sustainable model – the TARDiS road

- To achieve the original vision we are moving around the clock face
- Collaborating with academics to provide tailored valued services for different disciplines (needing extra functionality)
- Aided by a fast moving shared international movement

*All rising to great place
is by a winding stair*
Francis Bacon



Changing paradigms from 2000 - now

- Open Access little known
 - Open access publishing
 - Open access repositories
 - High level support
 - UK Science and Technology Committee
'Scientific Publications - free for all'
- Copyright transfer
 - License to Publish
 - Amendments
- Publishers
 - Changing policies
 - New publishing models
- Software, few options
 - Multiple, open source
- Funding Agency - no support
 - Declarations on Open Access signatories
 - Mandate deposit
 - Wellcome Trust
 - RCUK
 - NIH
- Authors non acceptance
 - Surveys - JISC/OSI Journal Authors Survey (3000 researchers)
 - 69% would deposit in IR if required by employer
 - 3% would not be prepared to do so
 - 66% thought archiving in IR important
 - 60% thought publishers should allow it

Of Course Researchers still have concerns :

Discipline differences
workload, status quo; content quality
control; authentication, versioning control
and of course Copyright

Benefits of Institutional Repositories

- Organization / University profile to Funding Agencies
- School and Discipline visibility
- Researcher profile
- Full text content
freely accessible,
more rapidly, easily
- Link to learning and teaching
- Links to projects and web pages

- Secure storage of publications for long term access to all research
 - including theses and dissertations, technical reports
- Stewardship of digital assets
- Demonstrates Interdisciplinary research
- Increased citations
 - Lawrence, Brody
- Research reporting

RAE management potential functionality – demo but planned enhancement

New JISC
Project to design
RAE module for use
within EPrints and
DSpace software
(Southampton and
Edinburgh)


Institutional
Repositories and
Research Assessment
(IRRA)

<http://irra.eprints.org/about.html>

JIBS 2005

Soton EPrints RAE Demo - Record for Sinha, Prof. Martin - Microsoft Internet Explorer

File Edit View Favorites Tools Help

 **e-Prints Soton**
University of Southampton e-Prints Service

[Home](#) | [About](#) | [Browse](#) | [Search](#)
[Register](#) | [User Area](#) | [Help](#)

PLEASE NOTE, this server is only a test installation to demonstrate and test RAE tools in eprints.
The real University of Southampton ePrints archive is at <http://eprints.soton.ac.uk/>

Record for Sinha, Prof. Martin

Please note, there have been some interface changes to this (very new) tool. The interface is not exactly as described in Pro easier and clearer now.

EPrint Records for RAE

Use the [My RAE Records Page](#) to view and edit the list of records in the eprint archive which will be considered for your part

Other information for RAE

Southampton Press Release 15 Dec 2004

University funded service managed by the
University Library

'We see our Institutional Repository as a key tool for the stewardship of the University's digital research assets,' said Professor Paul Curran, Deputy Vice-Chancellor of the University. 'It will provide greater access to our research, as well as offering a valuable mechanism for reporting and recording it.'

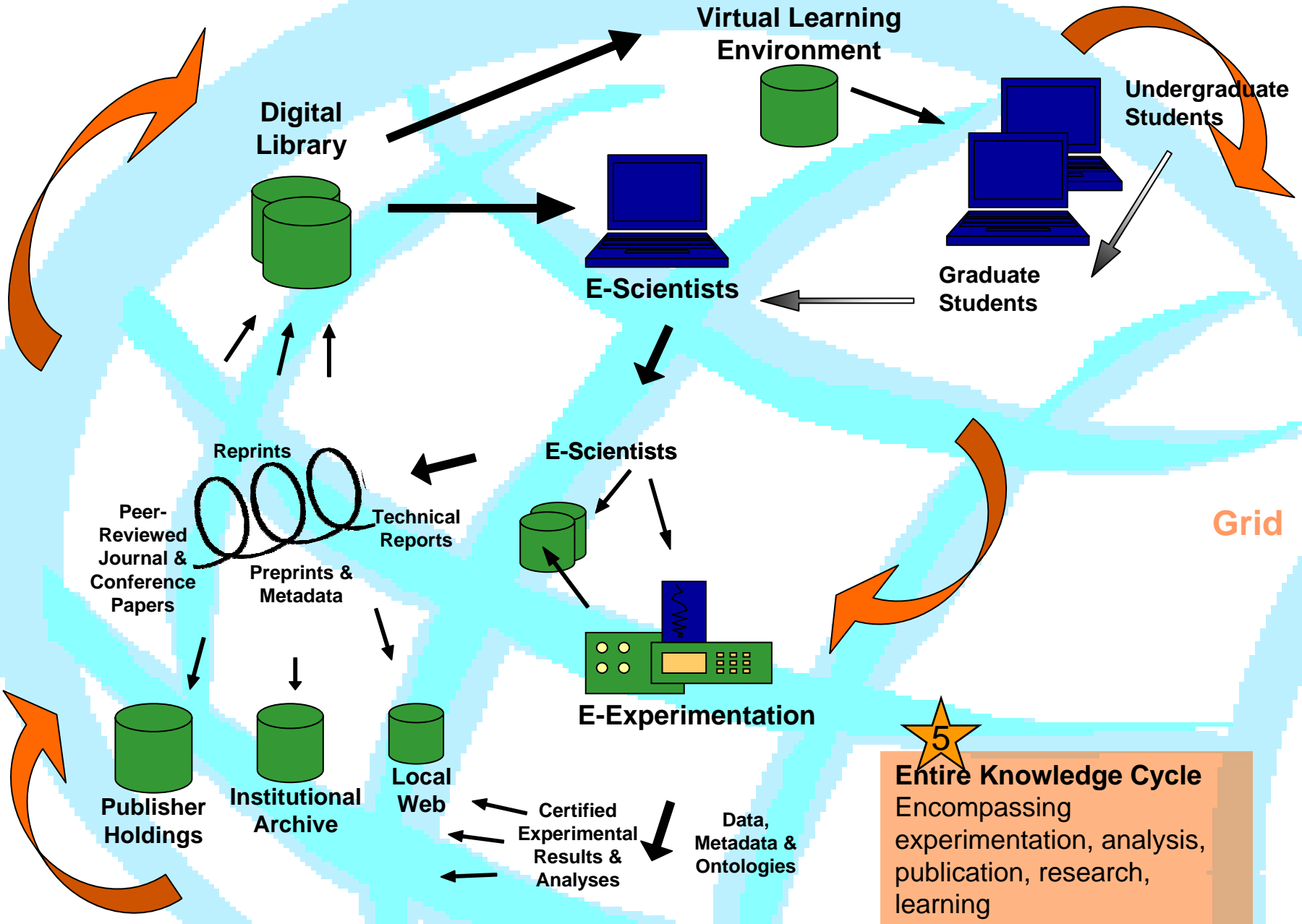



Repositories - a building block

- The Knowledge Cycle

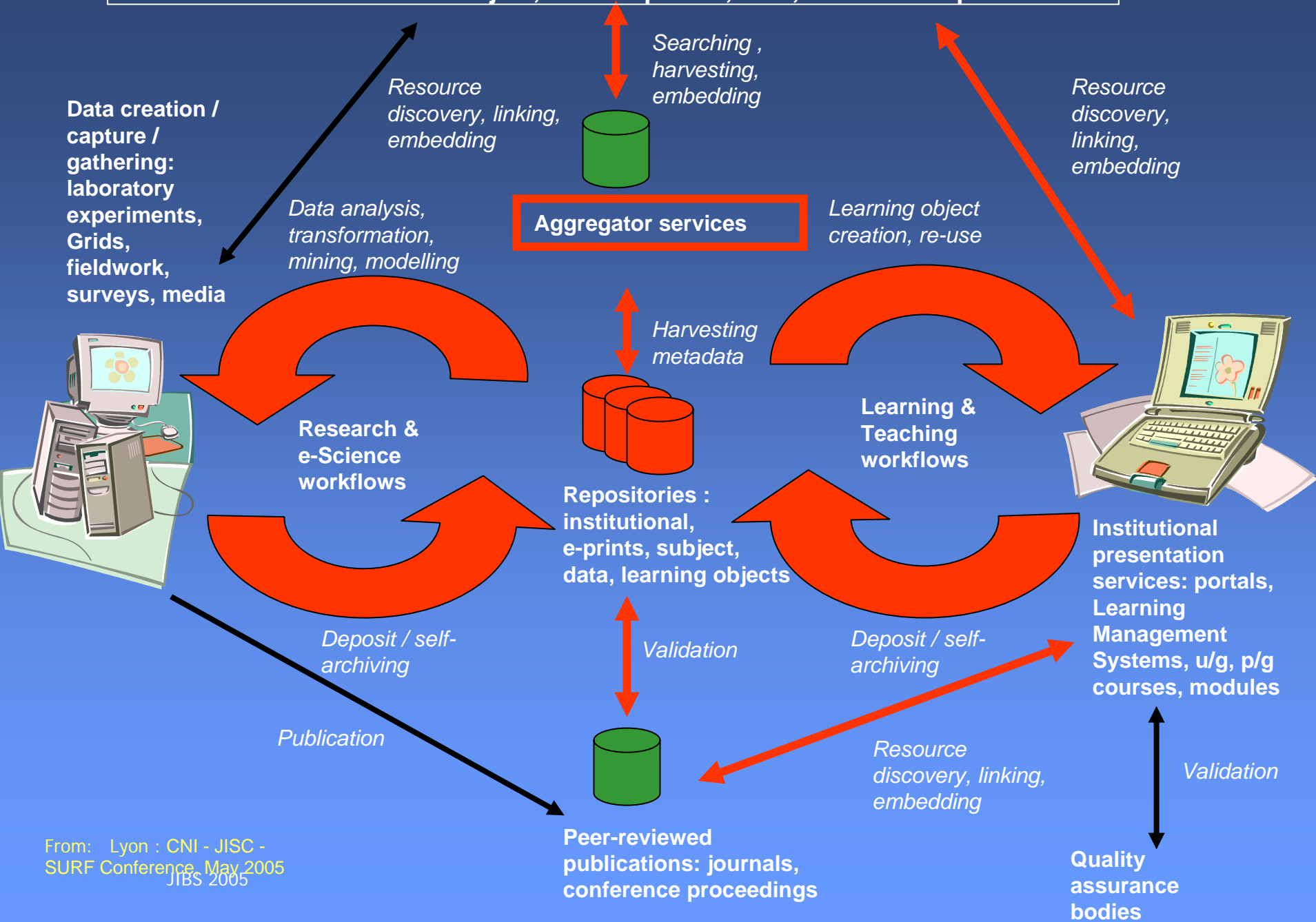
Encompassing experimentation, analysis, publication, research, learning

- Joined up research - Linking text and data
- An audit trail from whatever point of access




5
Entire Knowledge Cycle
 Encompassing
 experimentation, analysis,
 publication, research,
 learning

Presentation services: subject, media-specific, data, commercial portals



CLADDIER Project **

(Citation, Location And Deposition in Discipline and Institutional Repositories)

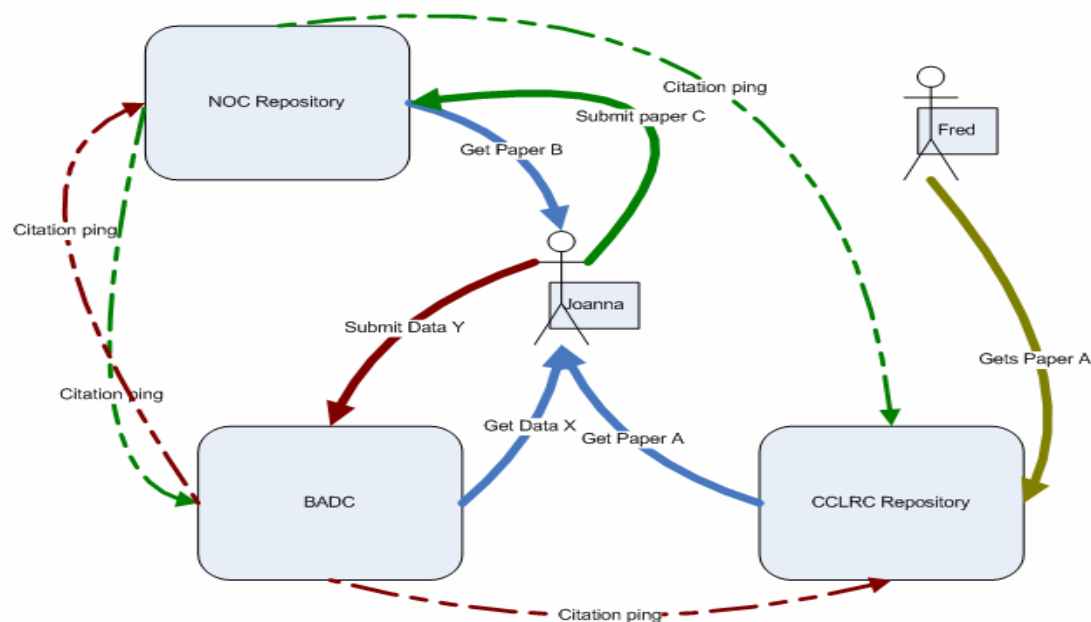
- The CLADDIER system will be a step on the road to a situation where (in this case, environmental) scientists will be able to move seamlessly from information discovery (location), through acquisition to deposition of new material, with all the digital objects correctly identified and cited. The lessons learned will be of applicability for the relationships between other discipline based repositories and institutional repositories.

**JISC Digital Repositories Programme 2005 -

CLADDIER

CLADDIER aims to build and deploy a demonstration system linking publications held in two institutional repositories (NOC and the CCLRC) with data holdings in the British Atmospheric Data Centre.

The use case below illustrates the proposed functionality.



1
Joanna gets data X and papers A and B for her research.

2
Joanna submits a paper C to NOC. The repository automatically checks and notifies the cited repositories with a "citation ping"

3
Joanna submits data Y to BADC. The data archive automatically checks and notifies the cited repositories with a "citation ping"

4
Fred gets the paper A from CCLRC. The paper "knows" it is cited in paper C and data Y.

Key outcomes


- Data is citeable in the same manner as papers
 - Papers are "aware" when they are cited

Persistent identifiers

Dataset Citations

Automated Linking both ways

A Plus for Researchers

- One outcome of CLADDIER Project
- 'pull' = Harvesting
- 'push' = CLADDIER outcome
 - Enable researcher to deposit in one repository and choose to upload (push) the metadata to another repository of choice (Institutional repository  repository of choice)
 - redundancy of records does it matter?

Building on TARDIS (the sequel(s))

- TARDIS completed its transition to invisibility early in 2005
 - **CLADDIER** (Citation, Location and Deposition in Discipline and Institutional Repositories) Linking e-Research. – partnering CCLRC, Reading, NERC
 - **PRESERV** (Preservation Services for EPrints) - partnering with National Archives File Format Registry (PRONOM) and the British Library, Oxford University (cooperation with SHERPA DP)
 - **GRADE** (Geospatial Repositories ...) – partnering EDINA et al

Back to the Future !!

Thank You

Pauline Simpson (ps@noc.soton.ac.uk)



**National Oceanography
Centre, Southampton**

UNIVERSITY OF SOUTHAMPTON AND
NATURAL ENVIRONMENT RESEARCH COUNCIL