



# Directions for Digital Repositories

Dr Leslie Carr

University of  
Southampton,

Web Science  
Institute

& EPrints  
Services

Texas Conference  
on  
**Digital Libraries**

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



# The Path Ahead

Institutional Repositories  
in Europe & the UK

Concluding Remarks

Current Research  
Information Systems

Open Access,  
Open Data  
and OER

# 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?

# Institutional Repositories in UK & Europe

- National differences
  - Centralised (FR)
  - Independent (UK)
  - Co-ordinated (NL)
- In UK, JISC funding synchronises community activity
  - VRE, e-Learning, Managing Research Data, Repository Enhancement, Community Content
- UK IRs typically chase embedded processes via subject liaison librarians
  - moving to “deskside support” model



# UK Institutional Repository Support

- Repository Support Project (RSP)
  - Provides training and support for repository staff
  - Helped set up repositories for institutions
- UK Council of Research Repositories (UKCoRR)
  - Independent organisation of repository managers
  - Unfunded
  - 197 members



# Open Access vs Research Assessment?

- In UK, high proportions of public-funded research
  - National research assessment processes
  - All universities required to account for research outcomes / quality
- Repositories find roles in cataloguing research outputs
  - Embeds them at the heart of institutional infrastructure
  - Diverts (temporarily) from Open Access goals

# Openness Agendas

- Open Access to Research Outputs
  - Supported by all Research Councils
  - Heavily promoted by JISC
- Open Research Data
  - Move from closed VRE / EScience environment
  - To open (able) data
- 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 Scientific Data

The figure is a screenshot of the eCrystals website. At the top, there is a navigation bar with links for Home, About, Browse by Year, Browse by People, Login, and Create Account. The University of Southampton logo is also present. The main content area displays the following information for a specific compound:

**Sample Originator:** Donated Sample<sup>a</sup>.

**Data Collection:** Susanne L. Huth<sup>b</sup> and Michael B. Hursthouse<sup>b</sup>

**Structure Determination:** Susanne L. Huth<sup>b</sup>.

Sample kindly provided by Maybridge<sup>a</sup>  
University of Southampton<sup>b</sup>

**Chemical Formula:** C<sub>10</sub>H<sub>9</sub>NOS

**InChI:** 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	C <sub>10</sub> H <sub>9</sub> NOS
Crystallisation Solvent	
Crystal morphology	Needle
Crystal system	Orthorhombic

**Available Files:**

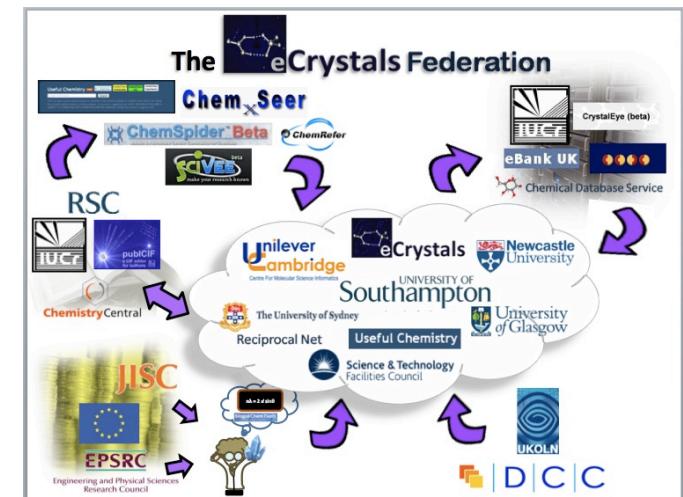
Final Result	<a href="#">2007may0014.cif</a>	11k
	<a href="#">2007may0014.cml</a>	3k
	<a href="#">2007may0014.fcf</a>	99k
Validation	<a href="#">2007may0014_checkcif.htm</a>	8k
Refinement	<a href="#">2007may0014.res</a>	4k
	<a href="#">2007may0014.xl.lst</a>	30k

**Chemical Structure:** A 3D ball-and-stick model of the compound (6-Thien-2-ylpyrid-3-yl)methanol is shown. The structure consists of a thienyl group attached to a pyridyl ring, which is further attached to a methanol group. The model is displayed in a Jmol viewer.

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

JISC EBank project (2003)

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



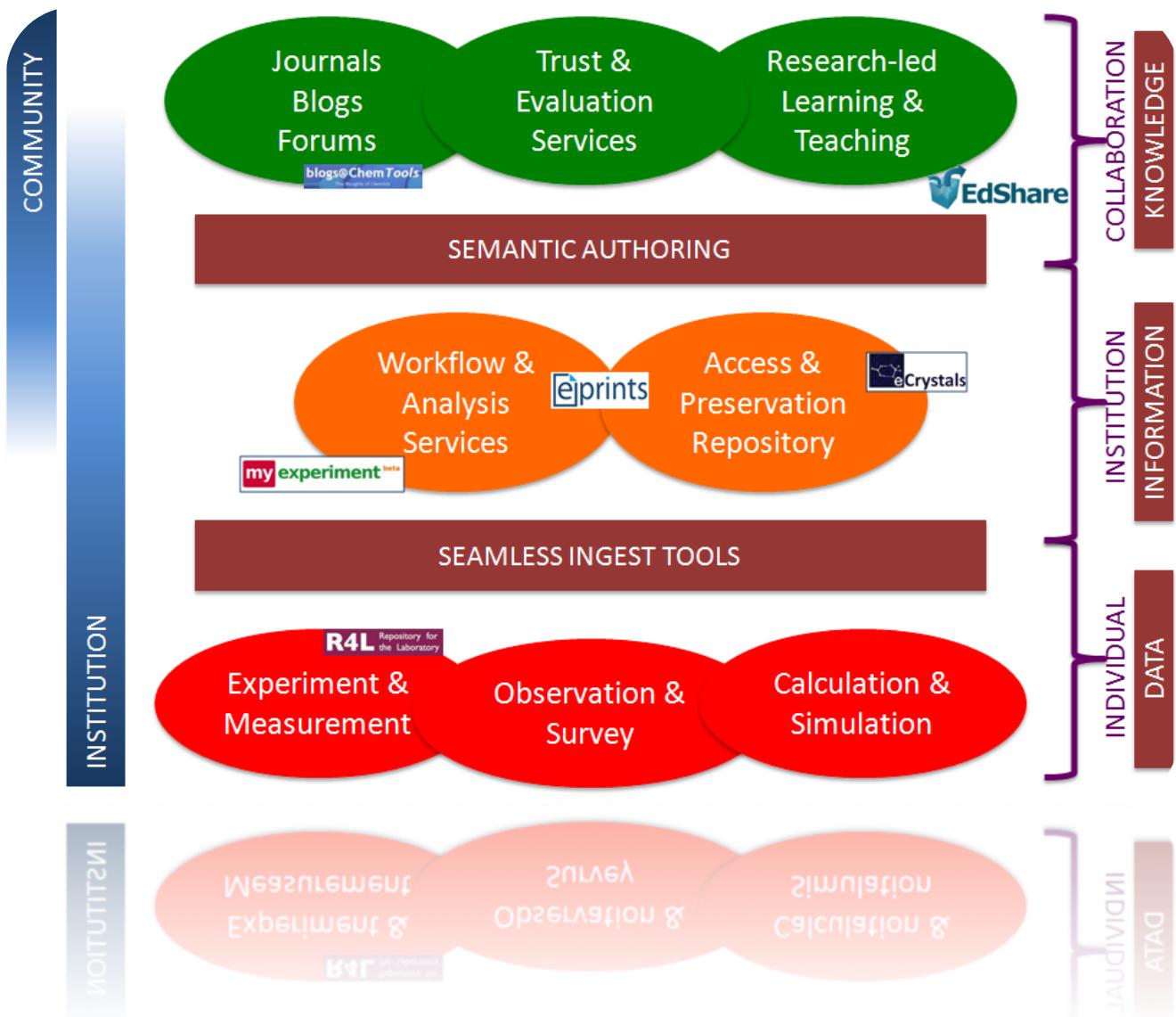
# Open Research

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

The screenshot shows the myExperiment website interface. At the top, there is a navigation bar with links for 'About', 'Mailing List', 'Publications', 'Log In', 'Register', 'Give us feedback', and 'Invite'. Below the navigation bar, the 'Workflows' tab is selected. The main content area displays a workflow entry titled 'Workflow Entry: Lymphoma type prediction based on microarray data'. The entry was created on 23/04/09 at 17:10:37 and last updated on 11/05/10 at 19:04:33. It has 3 credits, 11 tags, and 0 reviews. The workflow type is Taverna 2, and the original uploader is Wei Tan. The workflow diagram is a complex graph of nodes and connections, with a preview image showing the flow. The right sidebar contains sections for 'Workflow Type' (Taverna 2), 'Original Uploader' (Wei Tan), 'License' (CC BY SA), 'Credits' (3), 'Attributions' (0), 'Tags' (11), and 'Shared with Groups' (0). The sidebar also includes a 'Log in / Register' form and links for 'Need an account?' and 'Forgot Password?'. Popular tags listed on the sidebar include benchmarks, bioinformatics, BLAST, d-grid, data integration, ebi, example, graph, gwes, gworkflow, kegg, localworker, mygrid, ondex, ovtk, oxi, pathway, pathways, protein, pubmed, rdf, taverna, and text mining.

# Data Management

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



# Open Arts Data

## *Capturing artistic experimentation & performance vs capturing scientific experimentation*

# Open Educational Resources

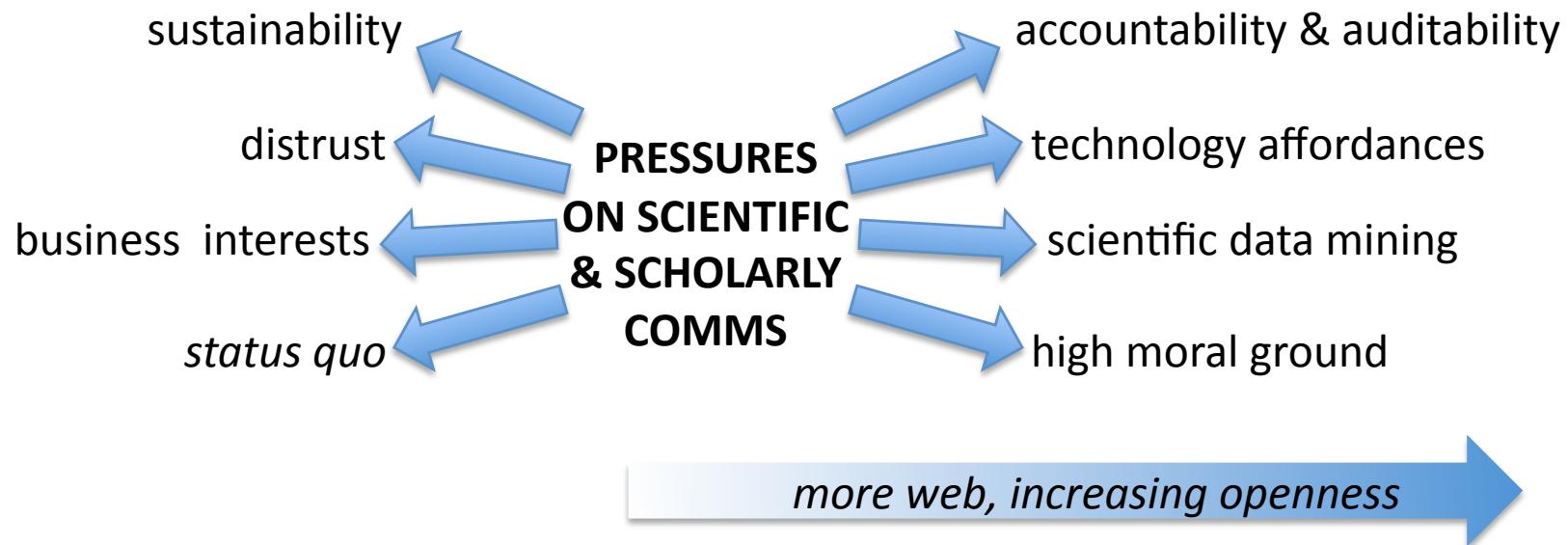
- JISC EdSpace
- IEEE LOM metadata vs
- Web 2, informally tagged
- Sufficient for discoverability and usefulness
- That includes *me!*

The screenshot shows a digital repository interface for 'EdShare' at the 'UNIVERSITY OF Southampton'. The top navigation bar includes 'Home', 'Browse', 'About', 'Help', 'Search...', and 'Advanced search'. Below the navigation is a secondary menu with 'Profile', 'Share manager', 'Bookmarks', 'Saved searches', and 'Logout'. The main content area displays a poster titled 'Posters for Web Science DTC Industrial Day'. The poster is titled 'What are the actual effects of Enterprise 2.0?' and includes sections on 'Origins of social software', 'Enterprise 2.0 Typology', 'Research methods', and 'Theoretical framework'. It also features a 'Hype and the question' section and a 'Research' section. The poster is attributed to 'Mark Schueler (mcs08@ecs.soton.ac.uk)'. Below the poster, there are download links for 'Mark Schueler poster.pptx' (Microsoft PowerPoint Download), 'Sarosh Khan poster.ppt' (Microsoft PowerPoint Download), 'Michael Yip poster.pdf' (PDF Download), and 'Mike Santer.pptx' (Microsoft PowerPoint Download). To the right of the poster, there is a 'Description' section, an 'Advice for reuse' section, a 'Share details' section, and a 'Toolbox' section with options like 'Edit item', 'Bookmark this item', 'Download as zip', and 'Use as template'. The 'Description' section notes that the posters were created by Web Science MSc and PhD students as a discussion point with representatives from the DTC Industrial Advisory Group. The 'Advice for reuse' section states that the posters are shared to facilitate ongoing discussions between students and the DTC faculty and industry.



JISC Jorum (full LOM) &  
JorumOpen (lightweight OA approach, DSpace)

# Directions for Researchers



- Open Access, Open Data, Open Educational Resources
  - Web Technology joins the High Moral Ground
- vs the established economic model for ensuring continuity of information production
  - trading of privately held information through subscription products such as journals or magazines

UEA appeared to support a culture of "resisting disclosure of information to climate change sceptics".  
... Prof. Phil Willis, described this as

The committee chairman, Phil Willis, described this as "reprehensible". Climate science must be transparent and irreproachable, the committee said.

# Climategate

from disciplines we all face

- Climate Research Unit at UEA refused to provide research data upon which its publications had been based

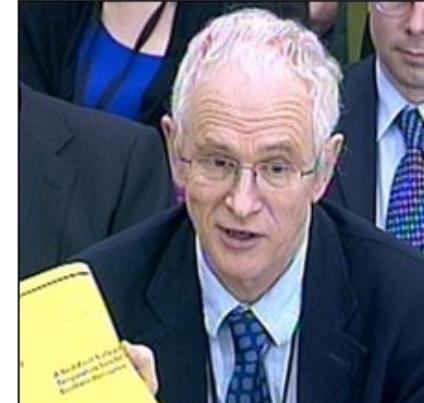
"When the prices to pay are so large, the knowledge on which these kinds of decisions are taken had better be right," the report stated.

## Climate science must be more open, say MPs

By Roger Harrabin  
Environment analyst, BBC News

**MPs investigating the climate change row at the UK's University of East Anglia (UEA) have demanded greater transparency from climate scientists.**

The Commons Science and Technology Committee criticised UEA authorities for failing to respond to requests for data from climate change sceptics.



BBC  
NEWS

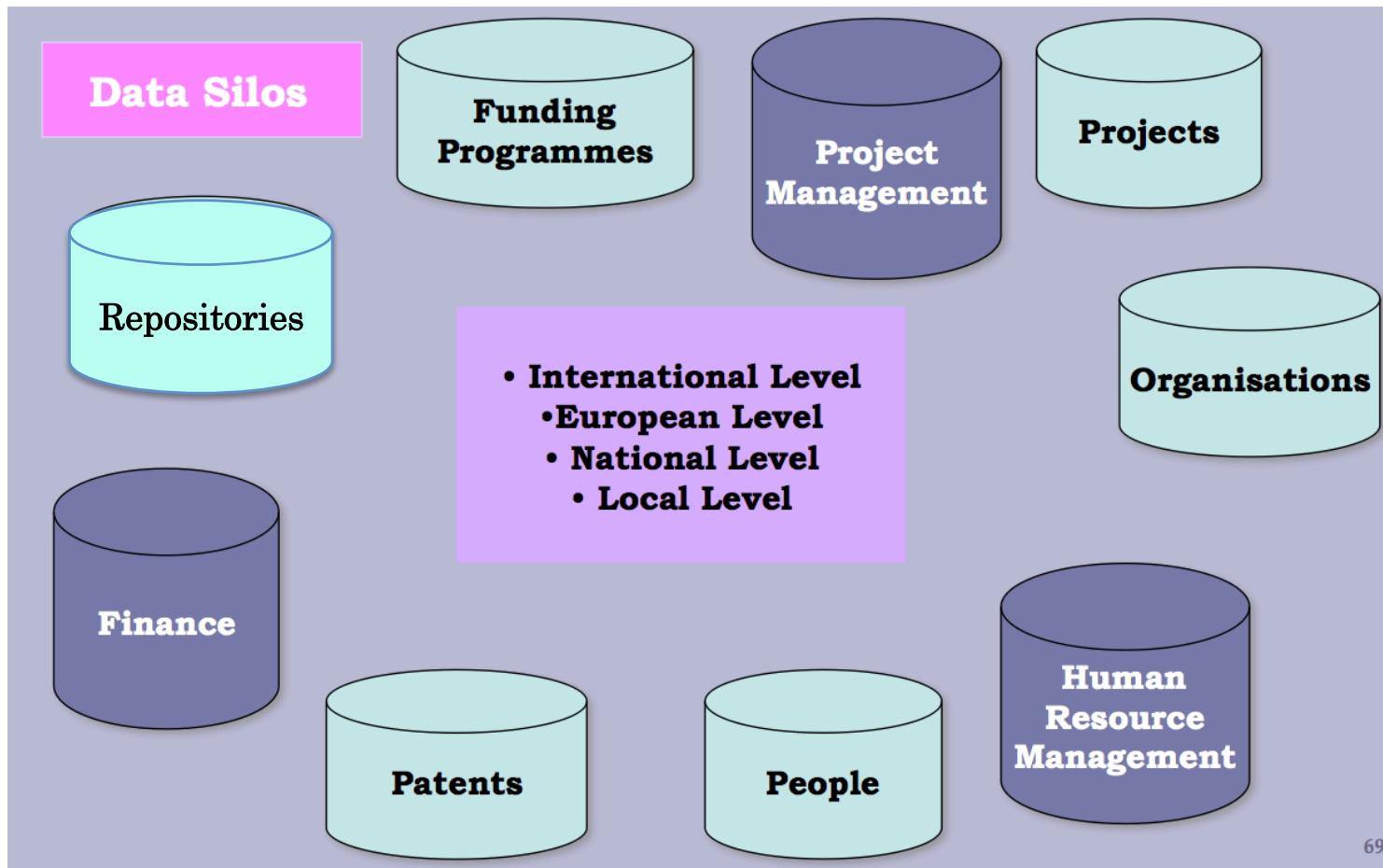
31 March  
2010

Roger Harrabin  
Environment analyst,  
BBC News

climate science to publish the raw data academic papers," its report said.

"We consider that climate scientists should take steps to make available all the data that support their work, including raw data... and full methodological workings, including the computer codes."

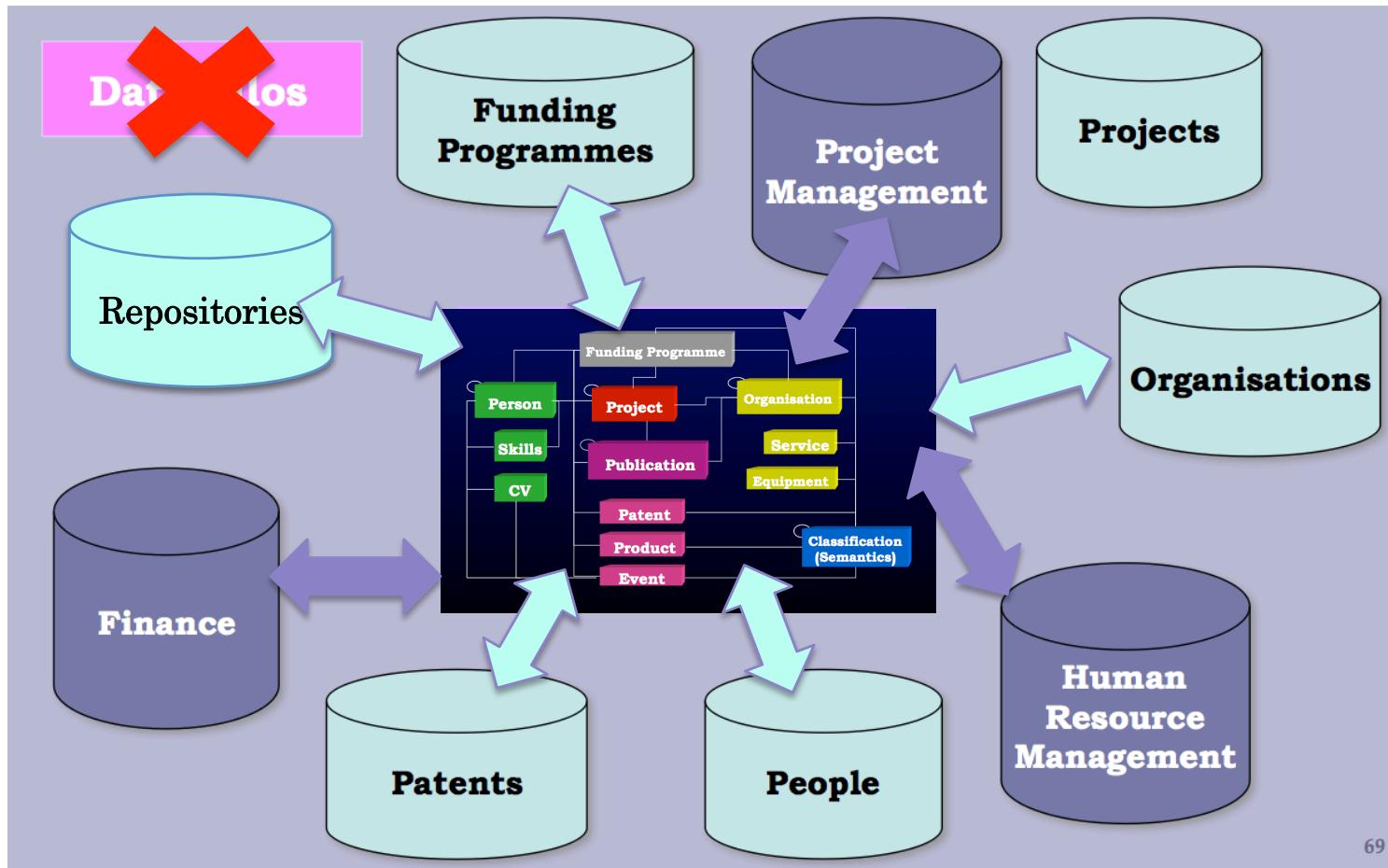
# CRIS: Current Research Info Systems



69

- EuroCRIS organisation promotes systems that deal with whole research infrastructure

# CRIS: Current Research Info Systems



- CERIF standard data schema interoperates between different systems

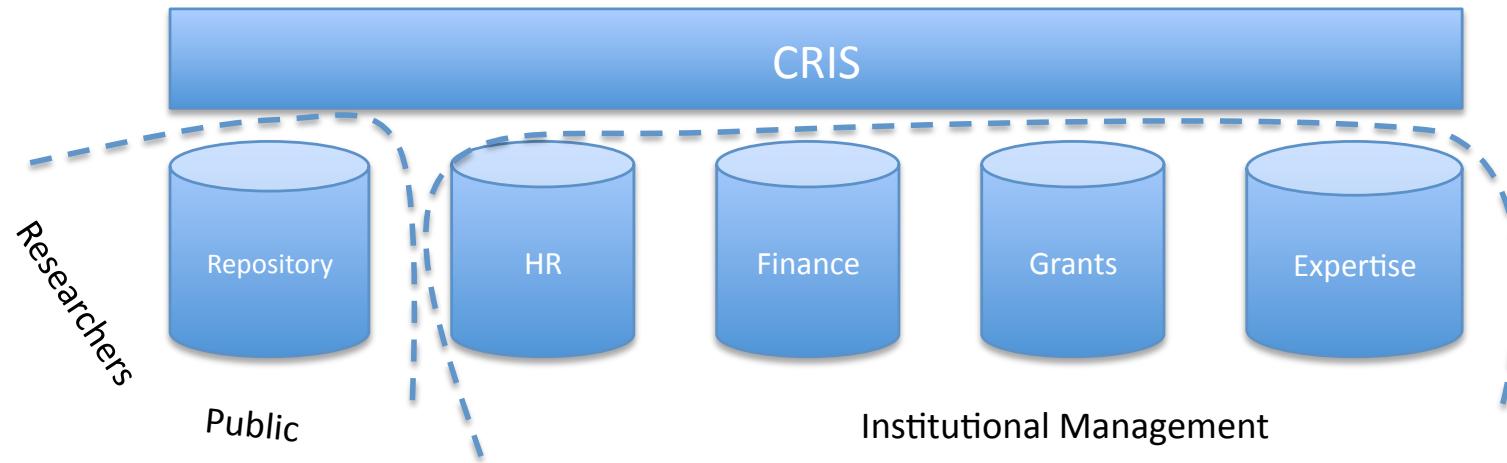
*Repositories should support CERIF*

# FRIS portal (NL)

- CRIS pulls together all research & admin information to create a unified view about a researcher's involvement
- Comes from admin
- Promoted by REF

The screenshot shows the FRIS research portal interface. At the top, there is a navigation bar with links for Help, Contact, About the Portal, News, and a search bar. Below the navigation bar, there is a blue header bar with links for Research projects, Organisations, Researchers (which is the active tab), Funding programmes, and Equipment. To the right of the header, there are language links for Nederlands and English. The main content area shows a researcher profile for 'Johan Lefevre' from K.U.Leuven, Department Biomedische Kinesiologie. It includes contact information like email (johan.lefevre@faber.kuleuven.be), address (UZ-Vesaliusinstituut-q, Minderbroedersstraat 12 blok q, 3000 Leuven), and a work address (http://cwisdb.kuleuven.be/persdb-bin/persdb?lang=N&cproep=persoon&fnam=00003378). Below the profile, there are tabs for Profile, Collaboration, Competence, and Map. The 'Collaboration' tab is active, showing a network graph of researchers and their affiliations. The graph includes nodes for Robert Vlietinck, Martine Thomis, Robert Fagard, Christophe Delecluse, Peter Hespel, Albrecht Claessens, Yves Vanden Auweele, and Bartholomeus Verreusel, all connected to Johan Lefevre. Some nodes also represent organizations like 'Departement Biomedische Kinesiologie' and 'Sportparticipatie en fysieke activiteit in Vlaanderen'. To the right of the graph, there are sections for 'Content', 'Researchers' (with a list of names like Kourosch Abbaspour Tehrani, Ahmed Abdelhamid, Ahmed Abdelkalek, and Ademola Aboye), 'Organisations' (with lists of department names), and 'Research projects' (with lists of project titles like 'Sportparticipatie en fysieke activiteit in Vlaanderen' and 'Steunpunt Cultuur, Jeugd en Sport'). At the bottom of the page, there is a footer with copyright information for the Flemish Government and a link to the Flemish Government website.

# CRIS Dichotomy

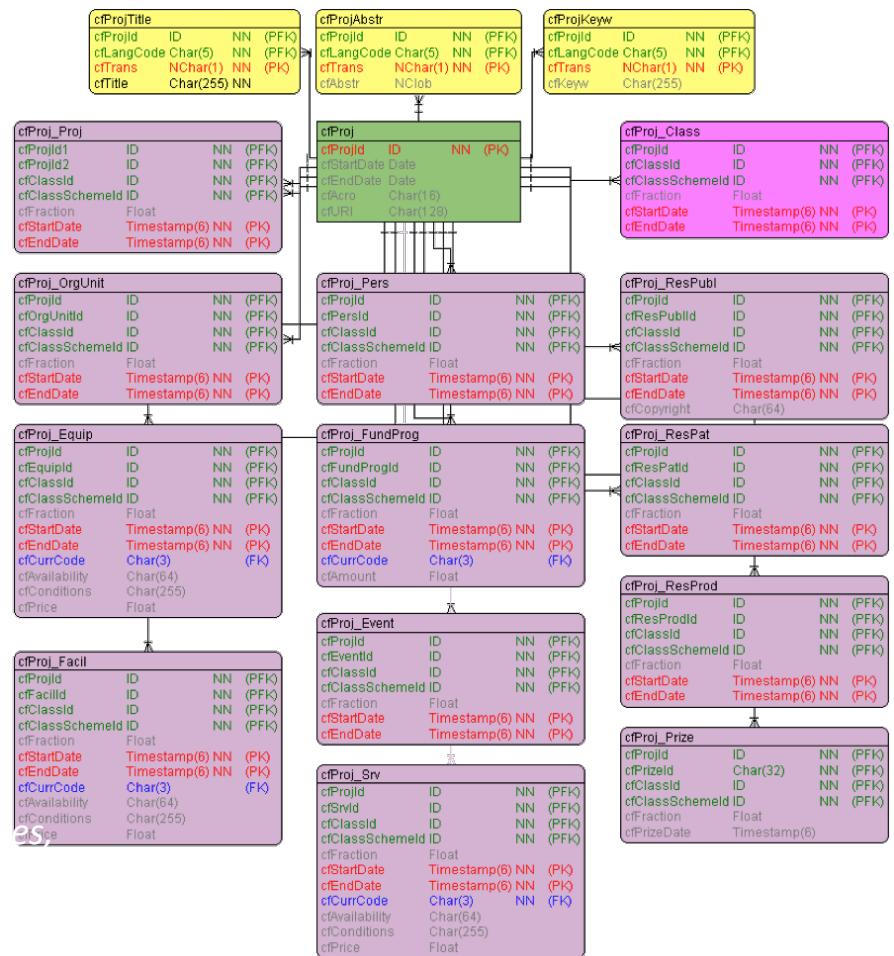


CRIS components attempt to provide service to  
*research management and administrators*

Repository attempts to reach researchers and  
public, and provide a service to management

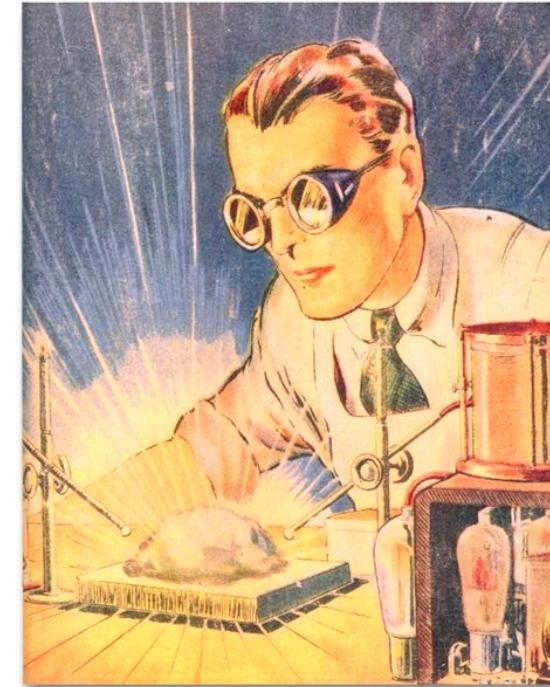
# Admin/CERIF View of A Project

- Grant ID / Funder / Amount
- Start date, End Date
- Investigators
- Budget breakdown



# Researchers’ View of A Project

- Project name
- Project aims / objectives
- Project Logo / Website / Blog
- Press releases
- News clippings
- Funder



# EdSpace

UNIVERSITY OF  
Southampton  
School of Electronics  
and Computer Science

Welcome to the EdTech project website!

is developing EdShare, the University of Southampton service to support sharing and collaboration in teaching. EdSpace aims to change culture in the University of Southampton by facilitating all involved to share educational resources. Technically, this will be achieved by implementing social "Web 2.0" tools for a repository based on EPiServer. Culturally, this will be achieved by working with all the parties to discover how they will best benefit from such change.

 **EdShare**

EdSpace is a JISC-funded project within the Institutional Exemplars Programme. It is managed from within the School of Electronics and Computer Science in collaboration with other groups across the University of Southampton: iSolutions (formerly ISS), the Library, Student Services and the Learning and Teaching Enhancement Unit (LATEU).

Start date: 01/10/2007  
End date: 30/04/2009  
Funded by:  
 JISC

31/03/2009  
EdSpace Project JISC

27/02/2009  
EdShare homepage re-designed

New presentation of a range of features

## EdShare upgrade to v.1.3

## Bookmarking functionality

### and Notes

### More Updates

# Pulling Together

- JISC Open Impact
- Repository of High Impact Research
- Brings together staff, outputs, projects
  - Repository acts as a CRIS
  - Holds admin data
  - Collects outputs and evidence



DCS Academy of Computing

Login | Create Account

Home About Web Science

Description Business Description Journalism Description Political Description

Web Science is the study of the largest information construct on the planet (the Web), the engineering principles of the protocols and software that form it and the impact of universal information provision on society. The Web Science Research Initiative (WSRI) was set up in 2006 as the result of a Memorandum of Understanding between MIT CSAIL and University of Southampton, ECS. (The founding directors from ECS are Professors Wendy Hall, Nigel Shadbolt and Tim Berners-Lee.) The ambition was to coordinate and support the study of the decentralised information system that is the World Wide Web. Since the launch of this Initiative the concept of Web Science has been widely disseminated and is establishing itself as an important area of activity. WSRI 19s activities have focused on (i) articulating a research agenda for the broader scientific community, (ii) coordinating the development of Web Science educational material and curricula and (iii) engaging in thought leadership for this emerging field. In order to continue with these activities and in support of the global development of Web Science the Directors of WSRI have established a charitable body - the Web Science Trust (WST). Independent of the original founding institutions, the goal is to encourage the widest participation in the development of Web Science.

Key publications

Timeline of Web Science research activity  
Download (216Kb)

Key researchers

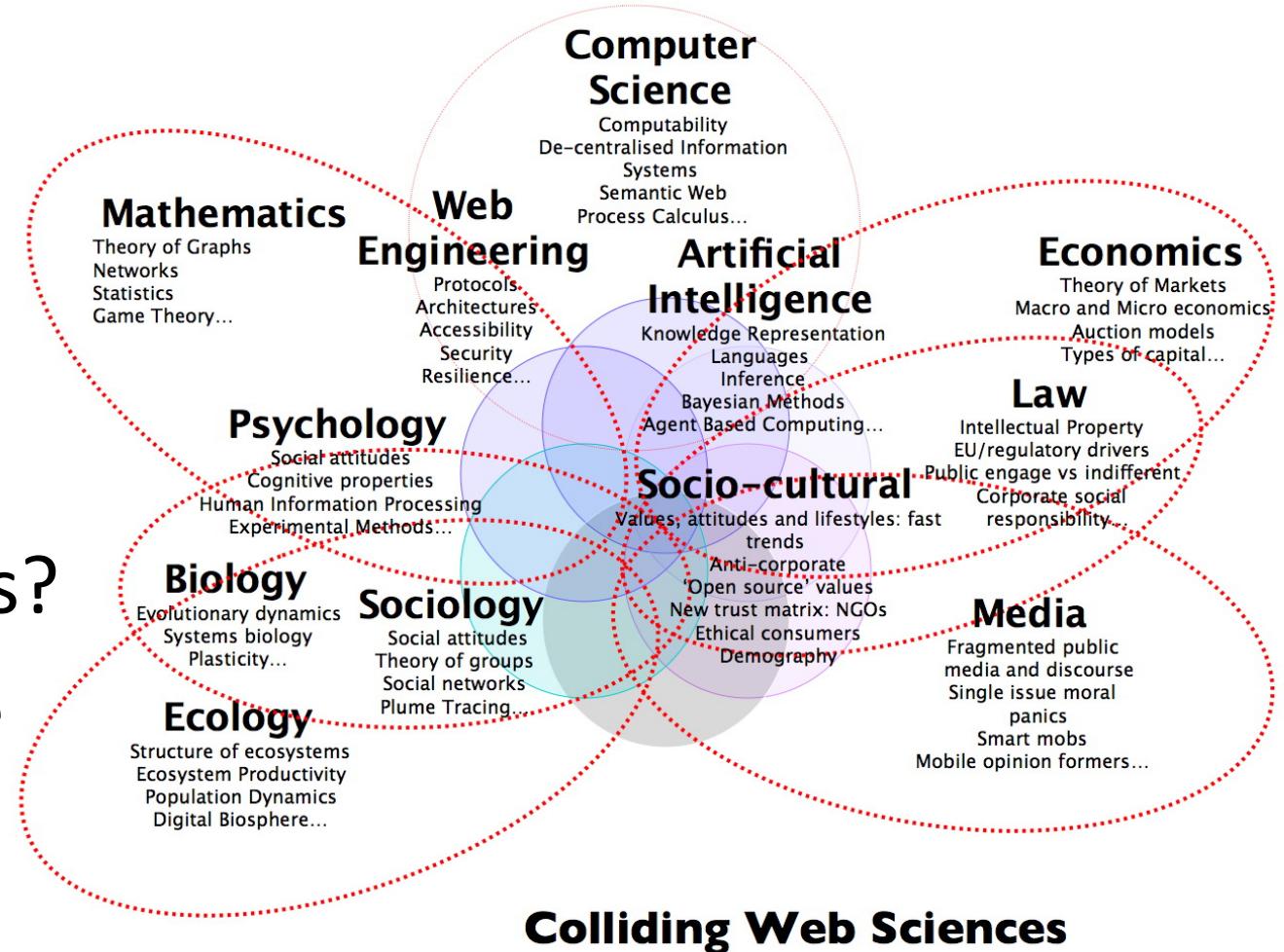
Wendy Hall, DBE, FREng, FRS is Professor of Computer Science at the University of Southampton, UK. The influence of her work has been significant in many areas including digital libraries & web science. She is a Founding Director of the Web Science Trust.  
[\[Further information\]](#)

Tim Berners-Lee is a Director and Trustee of the Web Science Trust (WST). He is a Director of the World Wide Web Foundation, started in 2008 to fund and coordinate efforts to further the potential of the Web to benefit humanity.  
[\[Further information\]](#)

Nigel Shadbolt is Professor of Artificial Intelligence (AI) and Deputy Head (Research) of the School of Electronics and Computer Science at the University of Southampton. is a Director and Trustee of the Web Science Trust.  
[\[Further information\]](#)

# In Conclusion: What Can We Do?

- What are digital libraries?
- What are repositories?
- What is the Web?



*Promote the study of the web and its role in society*



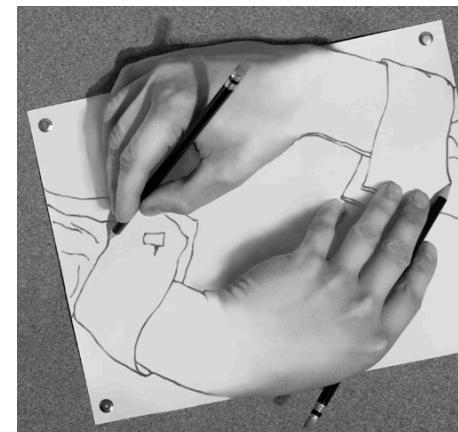
# Digital Libraries = Information Science + Social Network Science + Praxis

- The Web isn't a thing but an activity
  - the creation of a network of information by a network of individuals.
- The Web wasn't invented by Tim Berners-Lee,
  - it is being invented by all of us as we gradually adapt our tools and change our practice.

**The web both *shapes* and *is shaped* by society.**

There is a gap between the 'script' proposed by a technology and what it actually becomes in practice.

Co-constitution



# Librarians Mediating New Practice

Not just enforcing historic norms, but stimulating new practice to emerge

- Copyright
- Openness
- Intellectual Property
- Privacy
- Creativity
- Science 2.0

