



a centre of expertise in data curation and preservation

# Working in partnership with the eScience community – the Digital Curation Centre today and tomorrow

Graham Pryor

eScience Liaison

Digital Curation Centre, Edinburgh

Sheffield, 13<sup>th</sup> November 2008



This work is licensed under a Creative Commons Licence  
Attribution-ShareAlike 2.0





D | C | C

a centre of expertise in data curation and preservation

# Our objectives for today

- To learn more about eScience initiatives in Sheffield
- To tell you about some current DCC activities
- To identify (some of) your data issues and how we might help you
- To encourage and develop partnerships
- To identify next steps



Digital Curation Centre

[home](#) [contact us](#) [help](#)

Home

About the DCC

Newsroom

Events

Resource Centre

Tools & Standards

Welcome

The 4th International  
Digital Curation Conference

1-3 December 2008

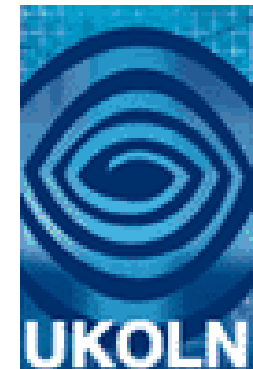
Edinburgh, Scotland



## UK Digital Curation Centre

<http://www.dcc.ac.uk/>

- Phase 2
- Community Development
- Curation Services
- Tools & Infrastructure
- Resources and Events
- Research Agenda
- SCARP Project
- Transition to Phase 3





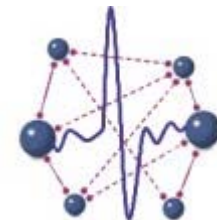
| D | C | C

a centre of expertise in data curation and preservation

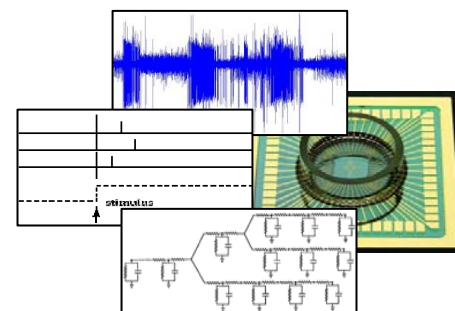
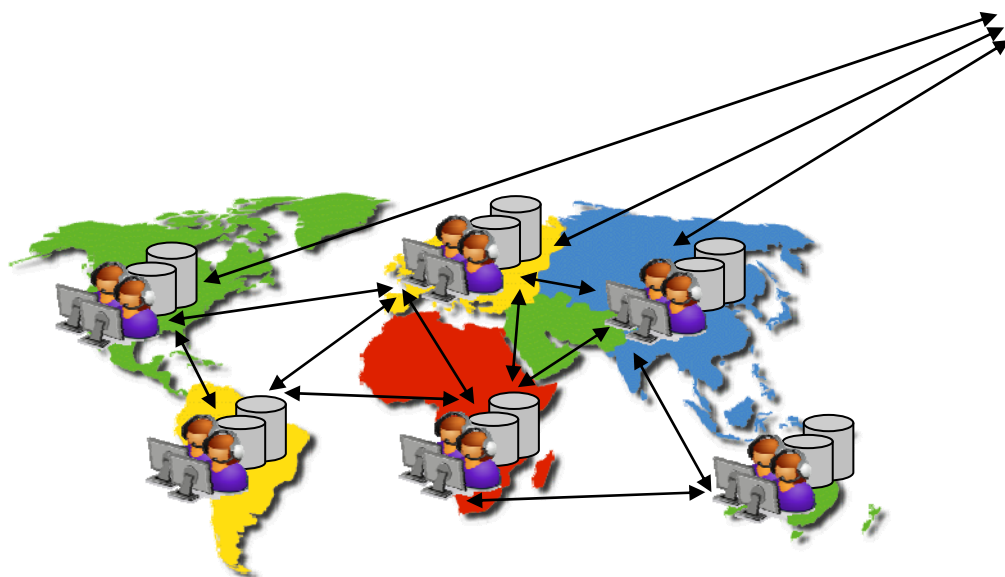
# Community development activities

- Data Centres support and “join-up”
  - Research Data Management Forum
- eScience Projects
  - Building closer links & partnerships –  
e.g. CARMEN

## CARMEN – <http://www.carmen.org.uk/>



Enabling sharing and collaborative exploitation of data, analysis code and expertise that are not physically collocated



Source: CARMEN SFN



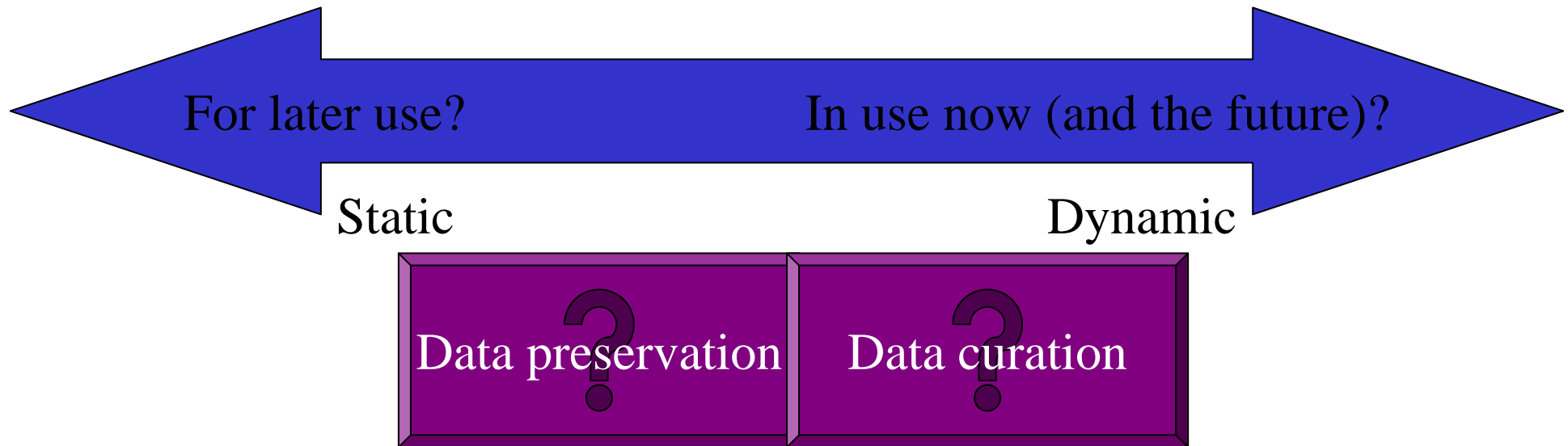
D | C | C

a centre of expertise in data curation and preservation

# Community development activities

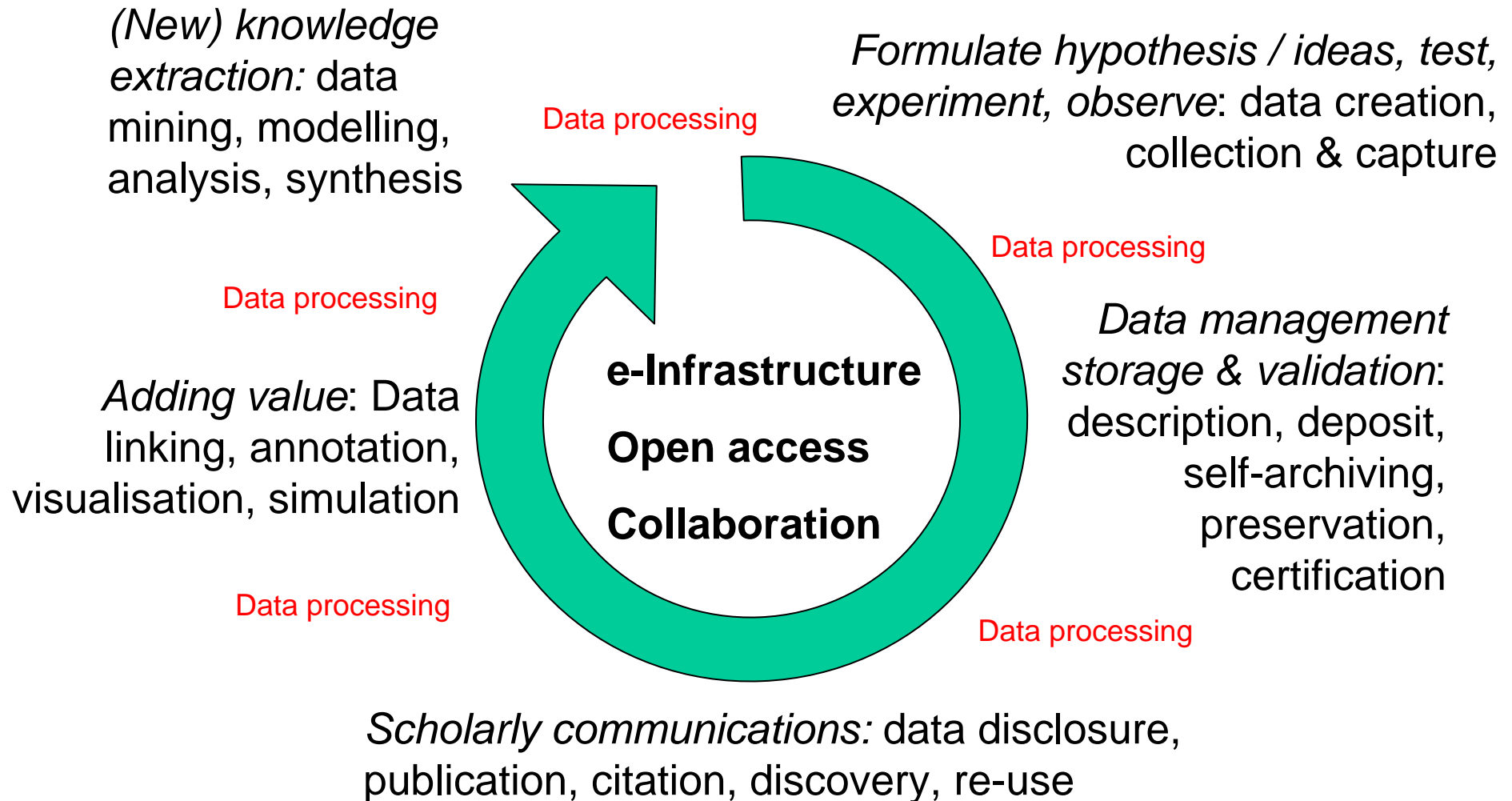
- Data Centres support and “join-up”
  - Research Data Management Forum
- eScience Projects
  - Building closer links & partnerships – e.g. CARMEN
- SCARP Project
  - Longitudinal, immersive case studies
  - Comparing individual discipline approaches to the creation, use and exploitation of data
  - Producing a register of best practice that crosses discipline and institutional borders

# Time-centric view of data curation?

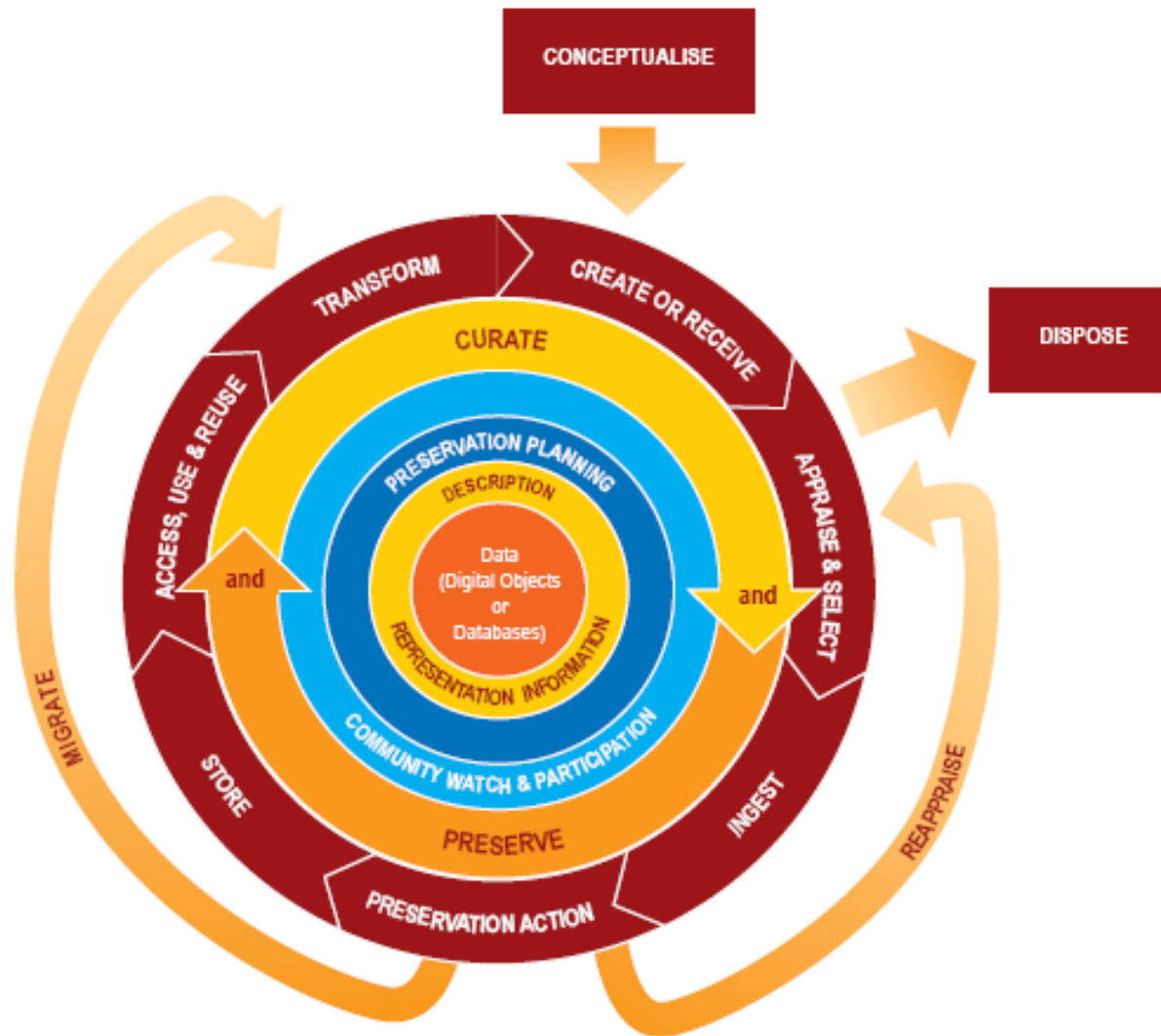


“maintaining and adding value to a trusted body of digital information for current and future use”

# Research lifecycle view of data curation?



# DCC curation lifecycle model



Designed by Sarah Higgins



D | C | C

a centre of expertise in data curation and preservation

## DCC Tools

- Diffuse Standards Frameworks -  
<http://www.dcc.ac.uk/diffuse/>
- Access to domain-specific information about the range of standards and specifications for curating and preserving access to digital materials
- Answers:
  - What standards should I be using?
  - When should I use them?



# D | C | C

a centre of expertise in data curation and preservation



## D | C | C

Digital Curation Centre

### DCC DIFFUSE Standards Frameworks

<http://www.dcc.ac.uk/diffuse/>

Sarah Higgins Sarah.Higgins@ed.ac.uk



Digital Curation Centre (DCC) DIFFUSE Standards Frameworks aims to provide access to domain specific information about the range of standards and specifications related to curating and preserving access to digital material. This will help curators, in a number of disciplines, identify which standards they should be using and at which stage in the lifecycle they should be implemented.

Standards Frameworks are sets of standards which can be used together to achieve effective implementation of all stages of the Digital Curation Lifecycle. A standards framework will ideally support interoperability, maximise accessibility, avoid vendor lock-in, provide architectural integrity, and help to ensure long-term preservation.

The forthcoming DCC Digital Curation Lifecycle Model will help to contextualise standards frameworks and relevant standards for a number of disciplines by presenting this information in a simple, visually descriptive format that can be utilised by a wide range of stakeholders.

Information provided for each standard and specification includes:

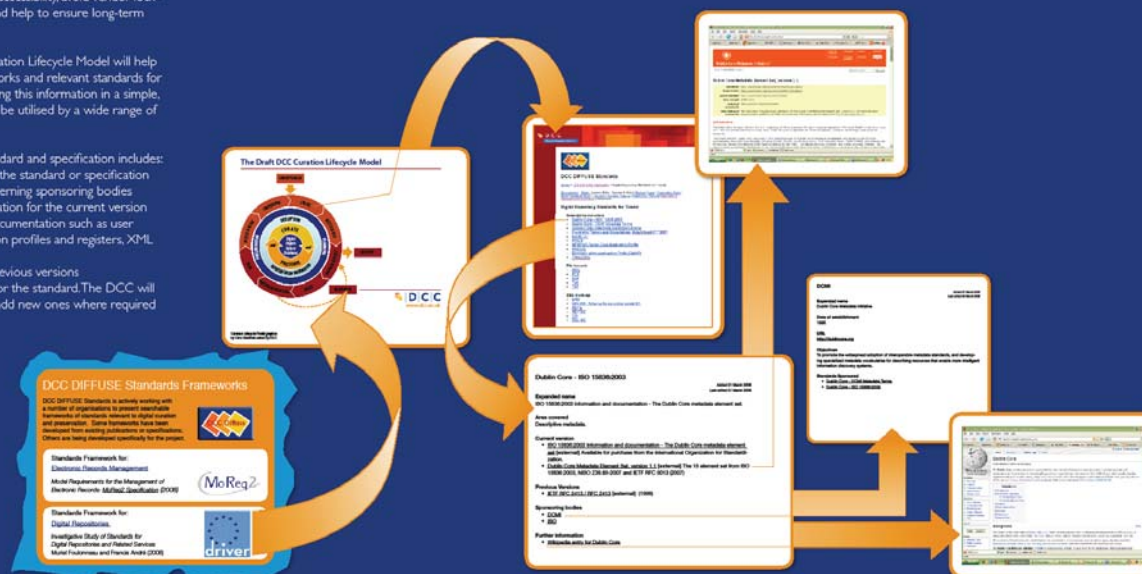
- A description of the scope of the standard or specification
- Links to database entries concerning sponsoring bodies
- Links to the official documentation for the current version
- Links to additional relevant documentation such as user guides, tutorials, implementation profiles and registers, XML DTD or Schema
- Links to documentation for previous versions
- Links to the Wikipedia entry for the standard. The DCC will curate Wikipedia entries and add new ones where required

#### Origins of DIFFUSE

The DIFFUSE Project was originally funded under the European Commission's Information Society Technologies (IST) Programme and ran from 1 February 2000 until 31 January 2003. Upon completion of the funding for this project the data they created was maintained as a valuable resource by TIEKE (the Finnish IT Development Centre), IC Focus and the SGML Centre. In 2005, the DCC secured permission to re-purpose the content of this database and are actively promoting the community-driven development of this newly dynamic resource.

#### Contribute to DCC DIFFUSE Standards Frameworks

The DCC is actively working with a number of organisations to present searchable frameworks of relevant standards. Disciplines currently represented include digital repositories, archives, records management, the museum sector and the geo-information sector. We are eager to collaborate with projects and organisations who would like to present an existing standards framework, or work with us to create one. Please send any suggestions, edits or related information to [info@dcc.ac.uk](mailto:info@dcc.ac.uk).





D | C | C

a centre of expertise in data curation and preservation

## DCC Tools

- **DRAMBORA – Digital Repository Audit Method Based on Risk Assessment -**  
<http://www.repositoryaudit.eu/>

The toolkit facilitates:

- Documentation of organisational and regulatory frameworks
- Identification of activities, assets and their owners
- Identification and assessment of risks associated with managing digital information
- Management of risks to ensure business continuity and future use of information



## DCC Tools

- **Data Audit Framework** - <http://www.data-audit.eu/>  
A process to identify, locate and assess the management of data assets:
- Planning the audit (the online tool enables scheduling and tracking)
- Identifying and classifying assets (an inventory is created, populated from an online questionnaire)
- Assessing the management of data assets (vital assets assessed in detail)
- Reporting and recommendations (including statistical analysis)

# Policy environment

<i>Data Policy</i>	BBSRC	MRC	Wellcome
Responsibility for data management /curation	With individual or institutional data custodians	With individual or institutional data custodians	Where possible, use recognised data repositories
Requirements for data management plan	Plan must be submitted with grant application Researchers responsible for data sharing	Plan must be submitted with grant application Access to MRC-funded data not to be restricted	Data plan accepted good practice; essential in cases of high data volume/perceived sharing benefit
Leverage for compliance with mandate	Funding included in project FEC Compliance monitored through institutional assessment	No funding released without approval of costed data management /sharing plan	Grants conditional on sharing plan and include data management cost Some equipment /database funding

## Policy environment

- Define current and future research data service needs
- Identify priorities for action
- Develop scenarios/options - from “do nothing” to a managed national service
- Develop business plan for preferred option(s), with costs/benefits
- Indicate scale of investment required and estimated ROI
- <http://www.ukrds.ac.uk/>



# International Journal of Digital Curation

[HOME](#) [ABOUT](#) [LOGIN](#) [REGISTER](#) [SEARCH](#) [CURRENT](#)

[ARCHIVES](#)

[OPEN JOURNAL  
SYSTEMS](#)

[Journal Help](#)

[Home](#) > [Current](#) > **Vol 1 (2006)**

## Vol 1 (2006)

[Table of Contents](#)

### Editorial

Editorial

*Liz Lyon, Chris Rusbridge*

[ABSTRACT](#) [PDF](#)

### Papers (Peer-reviewed)

Digital Curation for Science, Digital Libraries, and Individuals

*Neil Beagrie*

[ABSTRACT](#) [PDF](#)

Digital Curation, Copyright, and Academic Research

*Andrew Charlesworth*

[ABSTRACT](#) [PDF](#)

Scientific Publication Packages – A Selective Approach to the  
Communication and Archival of Scientific Output

*Jane Hunter*

[ABSTRACT](#) [PDF](#)

### Articles

A Grand Challenge: Immortal Information and Through-Life Knowledge  
Management (KIM)

*Alex Ball, Manjula Patel, Chris McMahon, Steve Culley, Stuart  
Green, John Clarkson*

[ABSTRACT](#) [PDF](#)

Joint US-UK Digital Preservation Workshop, Washington, D.C., May

[ABSTRACT](#) [PDF](#)

**Vol 2 No 2, 2007 published**

<http://www.ijdc.net/ijdc/issue/current>

### USER

Username

Password

Remember me

### JOURNAL CONTENT

Search

All

Browse

- [By Issue](#)
- [By Author](#)
- [Other Journals](#)

### INFORMATION

- [For Readers](#)
- [For Authors](#)
- [For Librarians](#)



D | C | C

a centre of expertise in data curation and preservation

# For discussion

- What are you currently doing to curate and preserve your data?
- What policies are in place?
- With which types of data, formats and metadata are you working?
- Where do you plan to store your data?
- Do you have a repository or content store?
- Do you have formal ingest processes?
- What specific data challenges do you have?