

DSpace Institutional Repositories and Digital Preservation

John MacColl

Digital Library Division

University of Edinburgh



What I will cover

- Why institutional repositories?
- DSpace and digital preservation
- A word about DIAS
- A word about the DCC
- Why libraries?





Edinburgh Research Archive



Edinburgh Research Archive >

Search ERA:

[Advanced Search](#)

Browse

- ▶ [Communities & Collections](#)
- ▶ [Titles](#)
- ▶ [Authors](#)
- ▶ [By Date](#)

Sign on to:

- ▶ [Receive email updates](#)
- ▶ [My ERA](#)
authorized users
- ▶ [Edit Profile](#)
- ▼ [Home](#)
- ▶ [Help](#)

Welcome to the Edinburgh Research Archive

Welcome to ERA, a digital repository of research output from The University of Edinburgh. Contained here are full-text digital Theses and Dissertations, book chapters, journal pre-prints and peer-reviewed journal reprints.

ERA is part of a world-wide network of Open Access Archives, which can be globally [cross-searched](#)

Search

Enter some text in the box below to search ERA.

Or go to the [Advanced Search](#)

Browse

- [Communities & Collections](#)
- [by Title](#)
- [by Author](#)
- [by Date](#)



Claims to allow digital preservation of locally generated resources

DSpace at MIT



[Home](#)

[Introducing DSpace](#)

[DSpace Technology](#)

[DSpace Federation](#)

[Implement DSpace](#)

[FAQ](#)

[News & Related Readings](#)

See DSpace in Action

[MIT](#)



About DSpace

DSpace is a groundbreaking digital library system to capture, store, index, **preserve**, and redistribute the intellectual output of a university's research faculty in digital formats.

Developed jointly by MIT Libraries and Hewlett-Packard (HP), DSpace is now freely available to research institutions world-wide as an open source system that can be customized and extended.

DSpace Announcements

- The DSpace Federation announces the creation of the [DSpace Federation Committers Group](#).
- MIT hosted a [DSpace user group meeting](#) on March 10-11, 2004.
- Join the [DSpace-General](#) and [DSpace-Tech mailing lists](#). Share your ideas and learn from other DSpace teams in these open forums.

DSpace Tech Notes

- Download the latest DSpace build from [Source Forge](#).
- Interested in [Contributing Code](#) to DSpace?

Resources

[Getting Started with DSpace](#)

[Communications Kit](#)

Feedback

[Are you running DSpace?](#)

[Discussion Forum](#)

[Mailing Lists](#)

MIT DSpace licence

Non-Exclusive Distribution License

In order for DSpace to reproduce, translate and distribute your submission worldwide your agreement to the following terms is necessary. Please take a moment to read the terms of this license, fill in the information requested (and sign and submit this license to Dspace at _____.)

By signing and submitting this license, you (the author(s) or copyright owner) grants to Massachusetts Institute of Technology (MIT) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video.

You agree that MIT may, without changing the content, translate the submission to any medium or format for the purpose of preservation.

You also agree that MIT may keep more than one copy of this submission for purposes of security, back-up and preservation.

You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright.

...



DSpace and digital preservation: two levels

DSpace identifies two levels of digital preservation: **bit preservation**, and **functional preservation**.

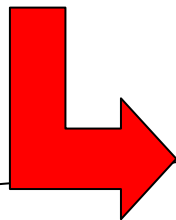
- **Bit preservation** ensures that a file remains exactly the same over time - not a single bit is changed - while the physical media evolve around it.
- **Functional preservation** goes further: the file does change over time so that the material continues to be immediately usable in the same way it was originally while the digital formats (and the physical media) evolve over time. Some file formats can be functionally preserved using straightforward format migration (e.g. TIFF images or XML documents). Other formats are proprietary, or for other reasons are much harder to preserve functionally.

...



DSpace and digital preservation: file format support

- **Supported:** We fully support the format and preserve it using either format migration or emulation techniques.
- **Known:** We can recognize the format, but cannot guarantee full support.
- **Unsupported:** We cannot recognize a format; these will be listed as "application/octet-stream", aka Unknown.

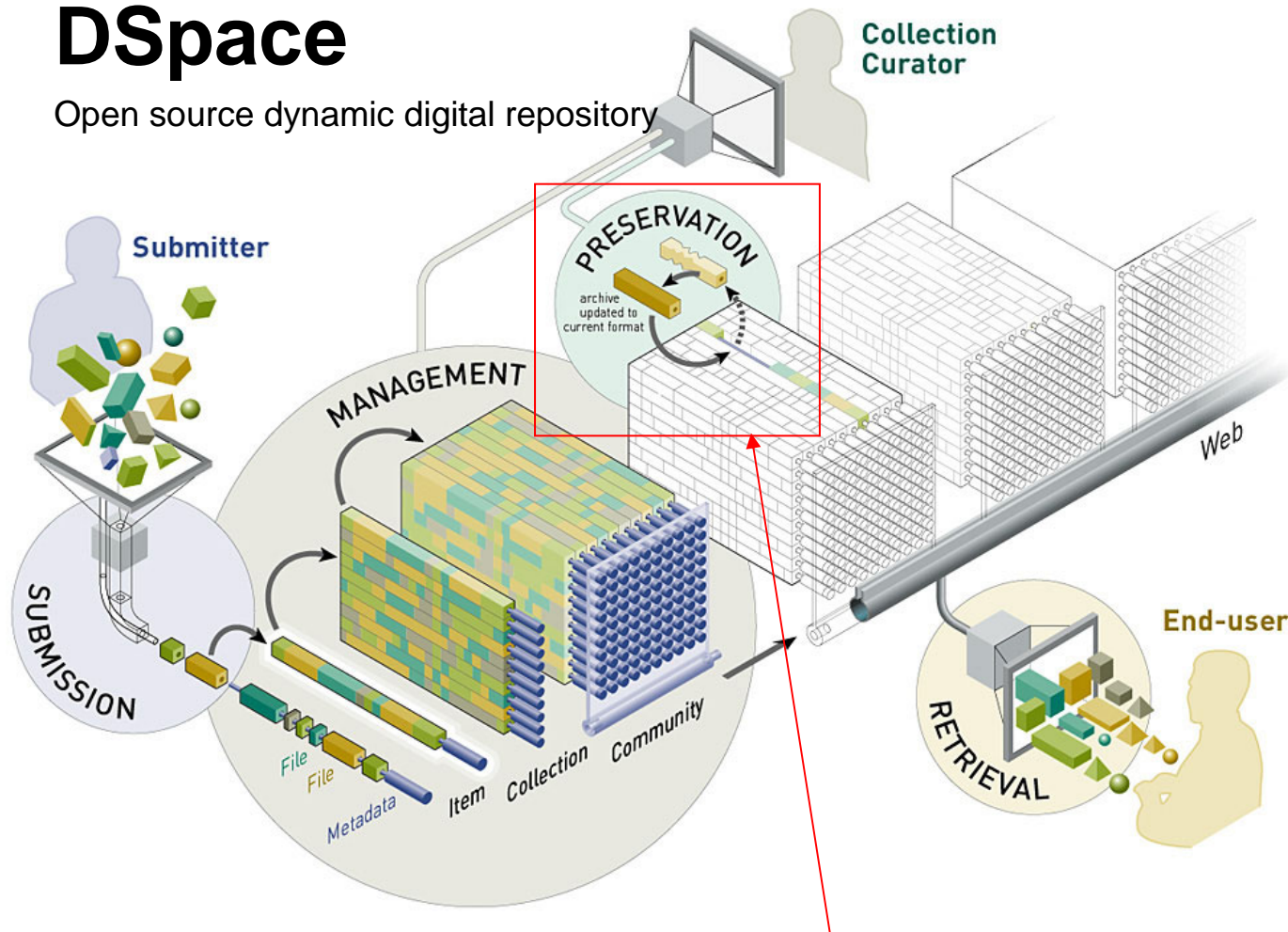


Bitstream preservation



DSpace

Open source dynamic digital repository



This slide is by MacKenzie Smith, Associate Director for Technology, MIT Libraries. Note the need for curation. The Preservation cycle shows format migration.



Paul Wheatley DSpace Futures Report (July 2003): Recommendations

- Allow migration on ingest or migration on request
- Accommodate loss quantification
- Record provenance
- Gather file format documentation (e.g. file specifications, descriptions and code samples)
- Create preservation action plans for particular formats




Edit View Favorites Tools Help

Back Search Favorites my Search More mywebsearch

Catalogi Collecties Inlichtingen Webexposities Organisatie Kenniscentrum

ONINKLIJKE BIBLIOTHEEK

mijn vraag ...



Zoek

in de catalogus
 op de website

zoek


uitgebreid zoeken >>>

Nieuws

- Overeenkomst KB en Oxford University Press over elektronische archivering
- Nieuwe KB-website op 19 oktober
- In de vitrine: Dichter des Vaderlands
- EU Conferentie
- Vernieuwde catalogus van kranten in de KB
- Week v.d. Geschiedenis
- Brinkman nieuwe voorzitter ABC
- 2 vacatures
- Nieuwsoverzicht

Uitgelicht

bladeren...



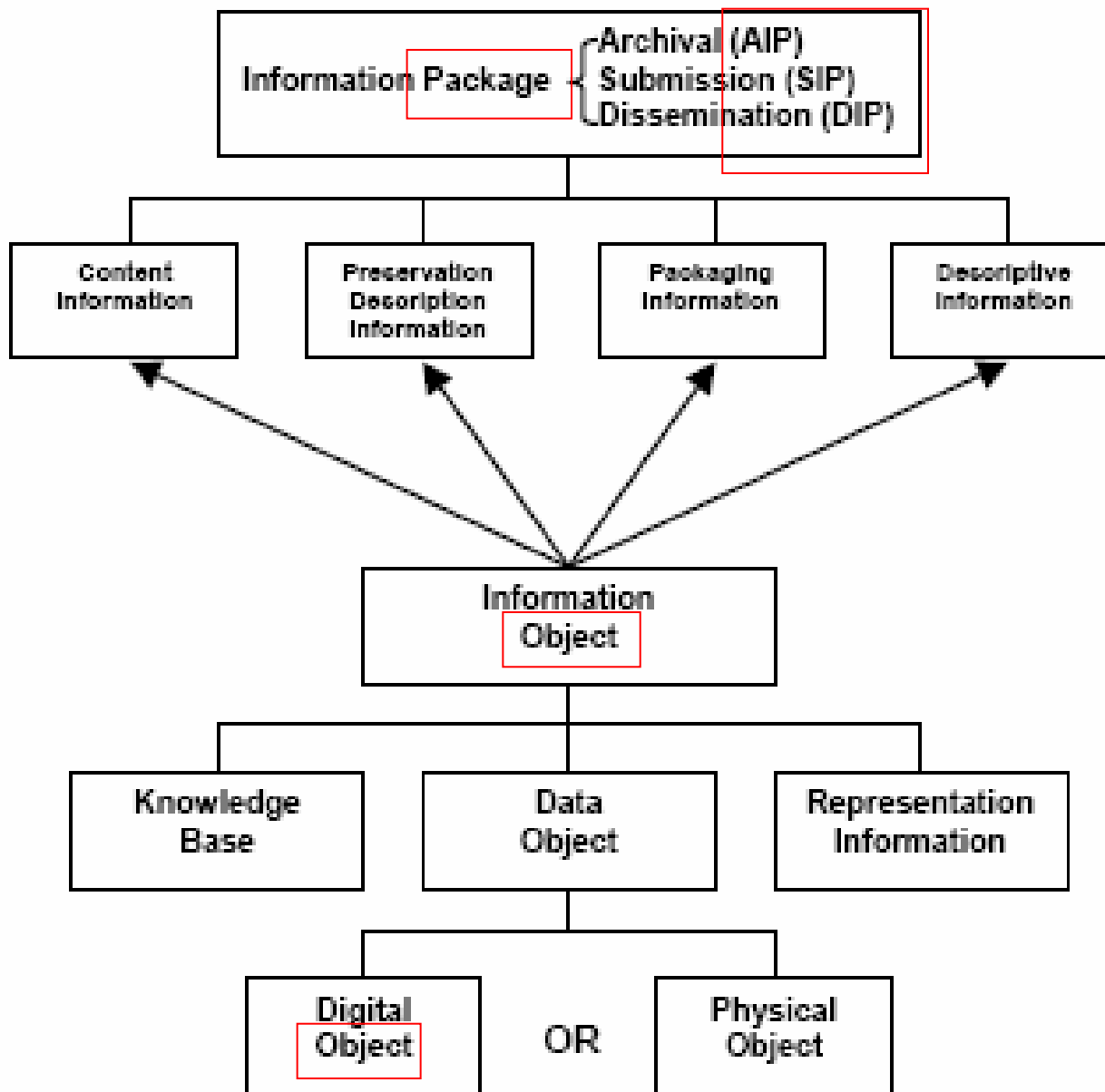
Stroetliederen

ATIONALE BIBLIOTHEEK VAN NEDERLAND

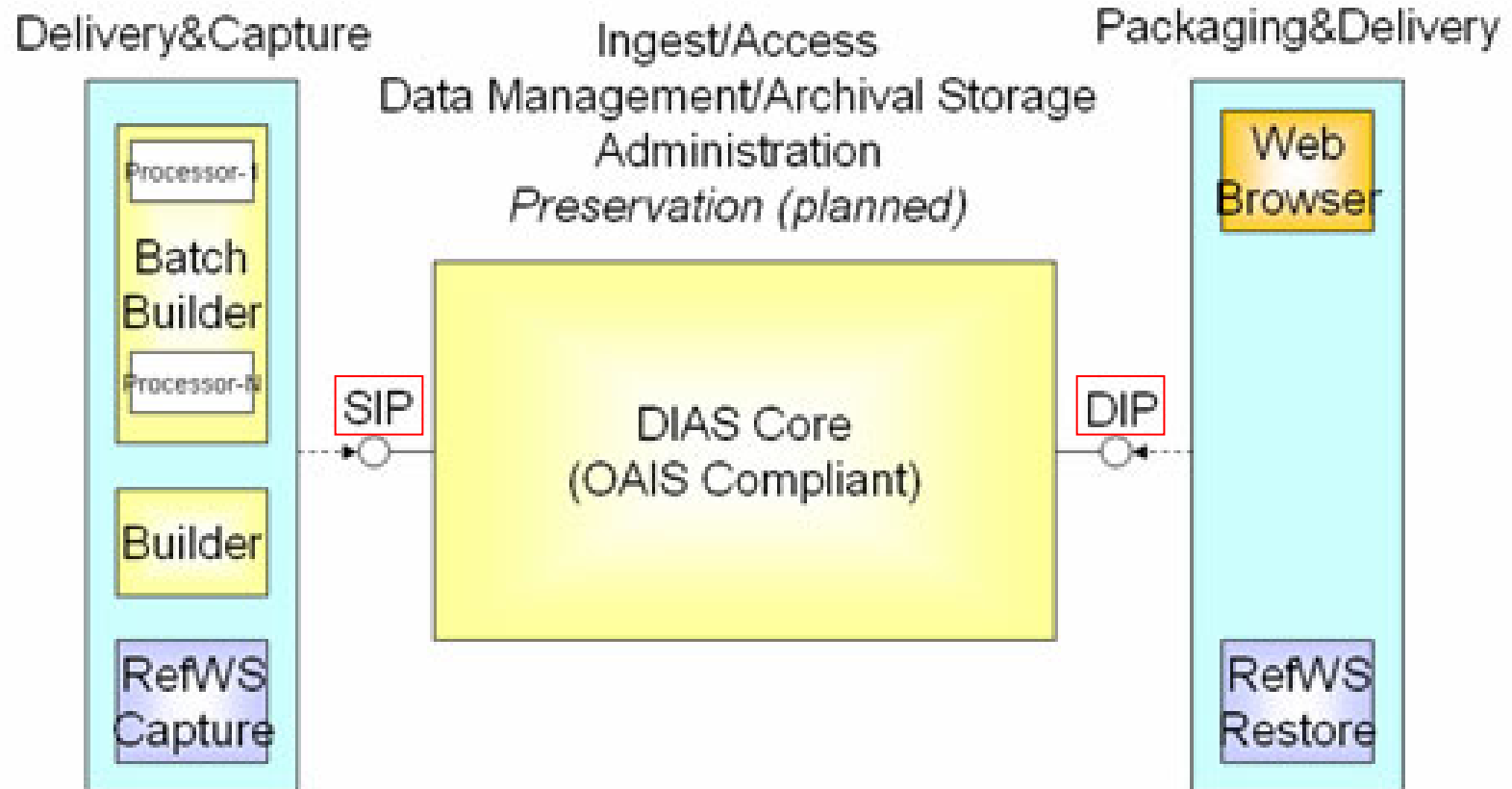
> ENGLISH

Internet

Figure 1: OAIS Information Model



Digital Information Archiving System (DIAS)



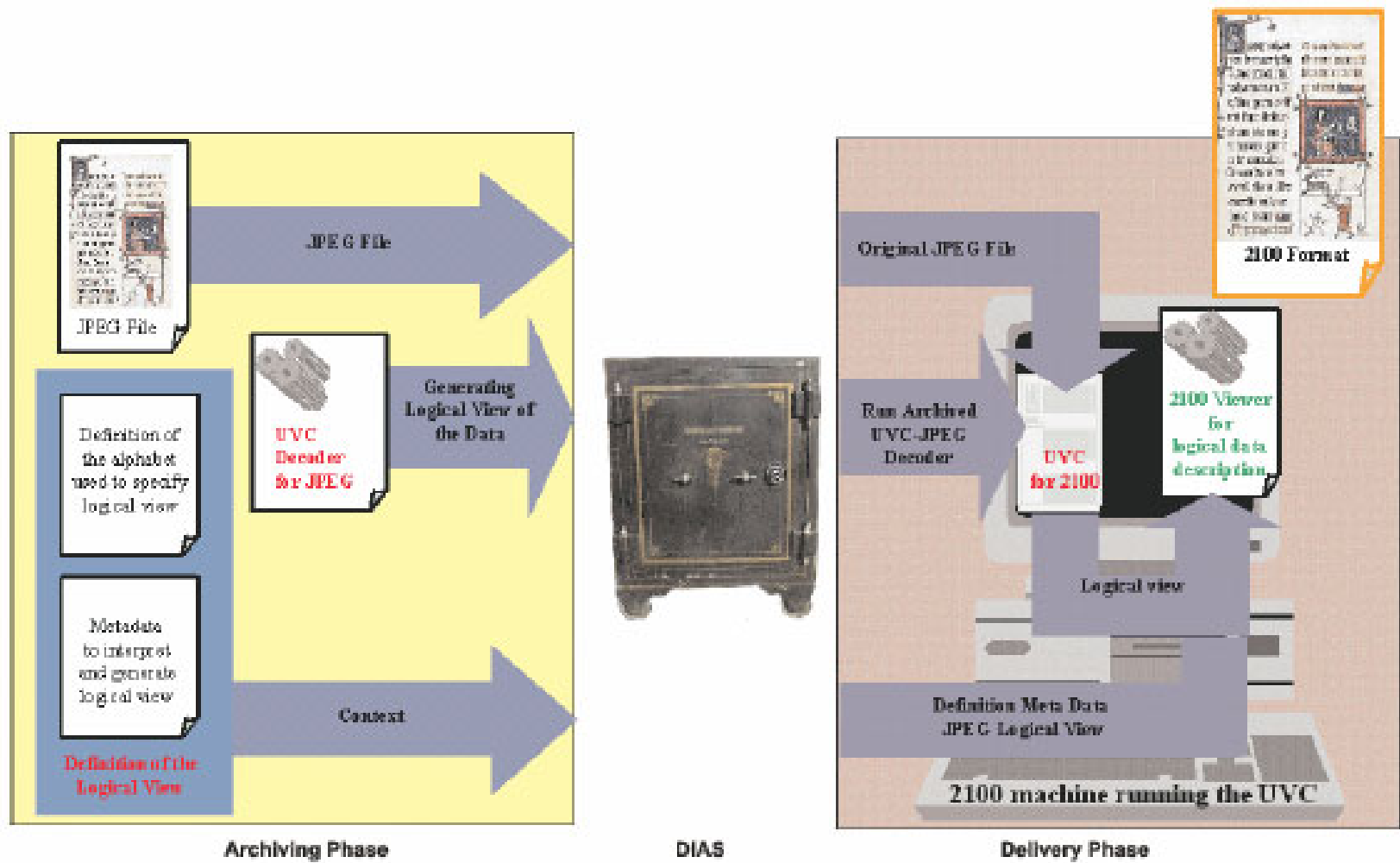
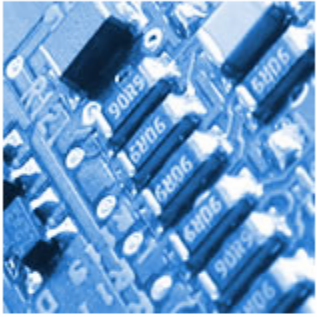


Figure 7.2 / Future rendering of digital objects in UVC-based Data Preservation

DCC Home

- Home >
- About the Digital Curation Centre >
- News and events >
- What is digital curation? >
- ACNDO (Associates Network) >
- DCC: preparing the launch >
- Search agenda >
- Contact information >

The JISC and the eScience Core Programme have entrusted the task of establishing the Digital Curation Centre (DCC) to a consortium comprising four partner institutions: the [University of Edinburgh](#) (lead partner) and the [University of Glasgow](#), which together host the NeSC; [UKOLN](#), at the University of Bath; the Council for the [Central Laboratory of the Research Councils](#) (which operate the Rutherford and Daresbury Laboratories).



Using the integrating ideas of collaboration, curation and continuing access for data held in institutional and national data repositories, we will be supporting a [Collaborative Associates Network of Data Organisations](#) in order that the DCC engages with and benefits from the progress being made by leaders within communities of practice across the wide range of scholarly and scientific disciplines, nationally and internationally.

We see our overriding purpose to be continuing improvement in the quality of data curation and digital preservation, recognising that data have importance as the evidential base for scholarly conclusions, and for the validation of those conclusions. In interests of securing consensus, we propose the term curation to cover the active management and appraisal of data over the corresponding lifecycle of scholarly and scientific interest: it is thus the key to reproducibility and reuse.

The DCC is not itself to be a digital repository, nor an attempt to impose policies and practices of one branch of scholarship upon another. Rather,

Libraries and the DCC

- Migration systems (e.g. TOM)
- Emulation systems (e.g. UVC)
- Representation Information systems (e.g. GDFR; PRONOM)
- Automated identification of file format
- Automatic extraction of file format metadata
- Automatic verification of format against an approved standard
- Automatic identification of dependencies
- Solutions for the content WE manage in our libraries
- There is a lot at stake ...



“I never think of the future. It comes soon enough.”

Albert Einstein, 1930.

