

An Introduction to the Digital Curation Centre

Chris Rusbridge

*Digital Curation Centre (DCC),
University of Edinburgh
c.rusbridge@ed.ac.uk*

Andrew McHugh

*Digital Curation Centre (DCC) &
HATII, University of Glasgow
a.mchugh@hatii.arts.gla.ac.uk*

Digital Curation¹ is the active management of data over the life-cycle of scholarly and scientific interest and is the key to reproducibility and reuse. Even before digital content is conceived, and after it has fulfilled its primary usefulness, digital curation activities must be undertaken and promoted if digital materials are to remain viable. Curation is not a box to be ticked on a manifest or a single process through which data passes. It is an ubiquitous endeavour that should characterise all interactions with and manipulations of digital content. Curation embraces and goes beyond that of enhanced present day reuse, and of archival responsibility, to embrace stewardship that adds value through the provision of context and linkage: placing emphasis on publishing data in ways that ease reuse and promoting accountability and integration.

The foundation of the Digital Curation Centre (DCC)² reflects the belief that long term stewardship of digital assets is the responsibility of everyone in the digital information value chain. The long term value of data rests in their potential as evidence, their reuse possibilities, and their role in facilitating compliance and in ameliorating risk. As scholarly research and scientific study become increasingly driven by the analysis of data, long term access to these data is crucial in enabling the verification of scientific discovery and providing a data platform for future research.

Only by promoting the ideas that underpin digital curation from the conception and curation of our digital assets until long after they have passed out of their primary usefulness can we claim to have succeeded. The Digital Curation Centre, through its organisation, emphases and practical activities closely reflects these ideals and it aims to catalyse action in innovative research, development, service delivery and outreach. Its primary aims are:

- to promote an understanding of the need for digital curation among the

1 This document is adapted from "*The Digital Curation Centre: A Vision for Digital Curation*" by Chris Rusbridge, Peter Burnhill, Seamus Ross, Peter Buneman, David Giarretta, Liz Lyon and Malcolm Atkinson, presented at *From Local to Global: Data Interoperability--Challenges and Technologies*, Mass Storage and Systems Technology Committee of the IEEE Computer Society, June 20-24 2005, Forte Village Resort, Sardinia, Italy

2 Digital Curation Centre, <http://www.dcc.ac.uk> [Accessed: 2 June 2005, 11:15]

communities of scientists and scholars;

- to provide services to facilitate digital curation;
- to share knowledge of digital curation among the many disciplines for which it is essential;
- to develop technology in support of digital curation; and,
- to conduct long term research into all aspects of digital curation

Led by the University of Edinburgh³, the DCC consortium includes HATII⁴ at the University of Glasgow⁵, UKOLN⁶ at the University of Bath⁷ and the Council for the Central Laboratory of the Research Councils⁸. Given the broadness and the pervasiveness of the digital curation challenge the core partners recognise that a sustainable contribution can only be made if widespread activity can be leveraged. To ensure that this happens the partners have established mechanisms to support the development of a network of associates.

Activities have been separated into four key task areas, with an umbrella management group overseeing and coordinating the work of each. The four main areas are Research, Development, Services and Outreach. By driving our activities from a solid research footing into innovative and service-led development work we can meet our numerous service objectives and ensure an effective and lucid outreach programme.

Within the research agenda, each topic involves fundamental problems, likely to generate high visibility outputs. Each will provide results that can be exploited within development activities and ultimately transformed into services. Current focusses include data integration techniques in the context of digital preservation metadata; effective data annotation techniques to facilitate subsequent searching, viewing and tracking across time, applications, researchers and migrations; appraisal for digital archiving; provenance and data quality; automatic extraction of semantic metadata; and legal issues. Current development effort is being concentrated mainly towards the creation of a representation information registry and repository, in order to document the structure and semantics of the ways in which digital data are stored, providing a method for accessing the content of digital objects. Representation information is not necessarily the original or official software access method or format specification,

3 University of Edinburgh, <http://www.ed.ac.uk> [Accessed: 2 June 2005, 11:15]

4 HATII (Humanities Advanced Technology and Information Institute) <http://www.hatii.arts.gla.ac.uk> [Accessed: 2 June 2005, 11:15]

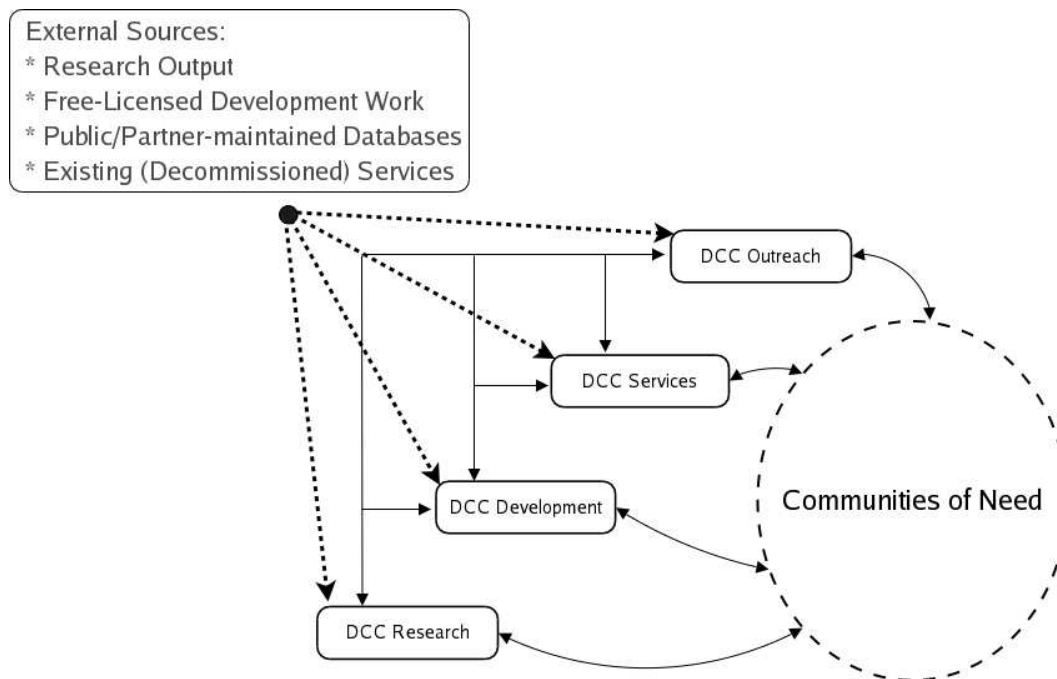
5 University of Glasgow, <http://www.gla.ac.uk> [Accessed: 2 June 2005, 11:15]

6 UKOLN, <http://www.ukoln.ac.uk> [Accessed: 2 June 2005, 11:15]

7 University of Bath, <http://www.bath.ac.uk> [Accessed: 2 June 2005, 11:15]

8 Council for the Central Laboratory of the Research Councils, <http://www.cclrc.ac.uk> [Accessed: 2 June 2005, 11:15]

but can take the form of anything that allows the information content of a digital object to be interpreted. Tools are being developed to reside atop the registry and repository infrastructure to facilitate its population, update and use.



Task Area Interrelationships

Services and outreach activities are regarded as vital if the DCC is to have the catalytic impact that's implicit in its remit. Current activities are wide ranging and directed at the very heart of community concerns and expectations. The DCC Advisory Service maintains a world class help desk facility, which offers a first point of contact for enquiries on a range of digital curation topics, contributing to and galvanised by an ongoing series of FAQs (*Frequently Asked Questions*). The DCC is also committed to the publication of a range of resources to further assist institutions, data centres and repositories with their digital curation efforts. At the forefront is the *Digital Curation Manual*, constructed from chapters contributed by international experts, with an editorial board consisting of leading researchers and practitioners in the area of digital curation. For many topics, a less in-depth insight is offered by the DCC briefing papers, designed to meet the needs of senior managers, offering quick and high level overviews of the topics that are explored in more detail in the Curation Manual. Additional DCC services include the development of audit and certification standards, in collaboration with the

Research Libraries Group⁹; the development of a repository of tools for digital curation; a preservation technology and standards watch; frequent case studies; and a series of regional and institutional site visits, which offer a useful opportunity to present details of the DCC's activities and establish a clearer understanding of community expectations. A further key service is the provision of a series of workshops and training events to reflect existing knowledge and practices, and to enable the community to work together to build new understanding in other areas. Forthcoming workshops focus on Persistent Identifiers¹⁰, Digital Repositories¹¹, and Costing Models¹² for Digital Curation. Outreach activities represent the forefront of the DCC's interactions with its stakeholder communities, and high visibility is essential to ensure that the Centre has sufficient community impact. User requirements gathering exercises; a 'virtual point of presence' Web portal; a high-quality peer reviewed *International Journal of Digital Curation*; and an Associates Network¹³ are all identified as key outreach deliverables. The latter aims to make the DCC partnership more pervasive; it welcomes and brings together prominent members from UK data creating and managing organisations, leading data curators overseas, supranational standards agencies, and representatives of UK industry and commerce involved in digital curation.

The viability of the DCC depends upon it gathering the right level of expertise, making that expertise available to the widest community, and demonstrating long-term commitment to the provision of research, development, services, and outreach. Encompassing insight, expertise and an acute awareness of the essential role of effective curation in all our digital activities, the DCC aims to be the embodiment of its vision for digital curation, and to succeed in its goals to be a standard bearer for best practice in an area that is relevant to every individual, institution and organisation that relies upon and uses digital information.

9 Research Libraries Group, <http://www.rlg.org> [Accessed: 2 June 2005, 11:15]

10 DCC Workshop on Persistent Identifiers, <http://www.dcc.ac.uk/piworkshop.html> [Accessed: 2 June 2005, 11:15]

11 DCC Workshop on Long Term Curation within Digital Repositories, <http://www.dcc.ac.uk/drworkshop.html> [Accessed: 2 June 2005, 11:15]

12 DCC and Digital Preservation Coalition joint workshop on Digital Curation Cost Models, <http://www.dcc.ac.uk/cmworkshop.html> [Accessed: 2 June 2005, 11:15]

13 Digital Curation Centre: Associates Network, <http://www.dcc.ac.uk/associates>
<http://www.dcc.ac.uk/associates/>