

 | D | C | C

New Organisational Structures Responding to New Challenges: The Digital Curation Centre in the UK

Public Lecture
Schweizerisches Bundesarchiv, Bern Switzerland
25 October 2004

Seamus Ross, DCC Associate Director
& Director HATII, University of Glasgow
In collaboration with: Peter Burnhill and Peter Buneman (Edinburgh), David Giarretta (CCLRC), Liz Lyon (UKOLN).


Funded by:  

 | D | C | C

UK Digital Curation Centre

- JISC Circular 6/03 called for bids in digital curation
- JISC and the e-Science Core Programme funding
 - for development, services, and outreach in digital curation
 - for a research programme
- Impetus to action
 - Growth in e-Science activity and data creation
 - Recognition that continuing access to digital materials crucial
- Goals of JISC 6/03 required ambitious and far-sighted vision—research, action, and support
 - across disciplines (research councils), HE, FE and beyond
 - across data types

2

 | D | C | C

©DCC. S Ross et al, HATII, UofGlasgow.

DCC Purpose

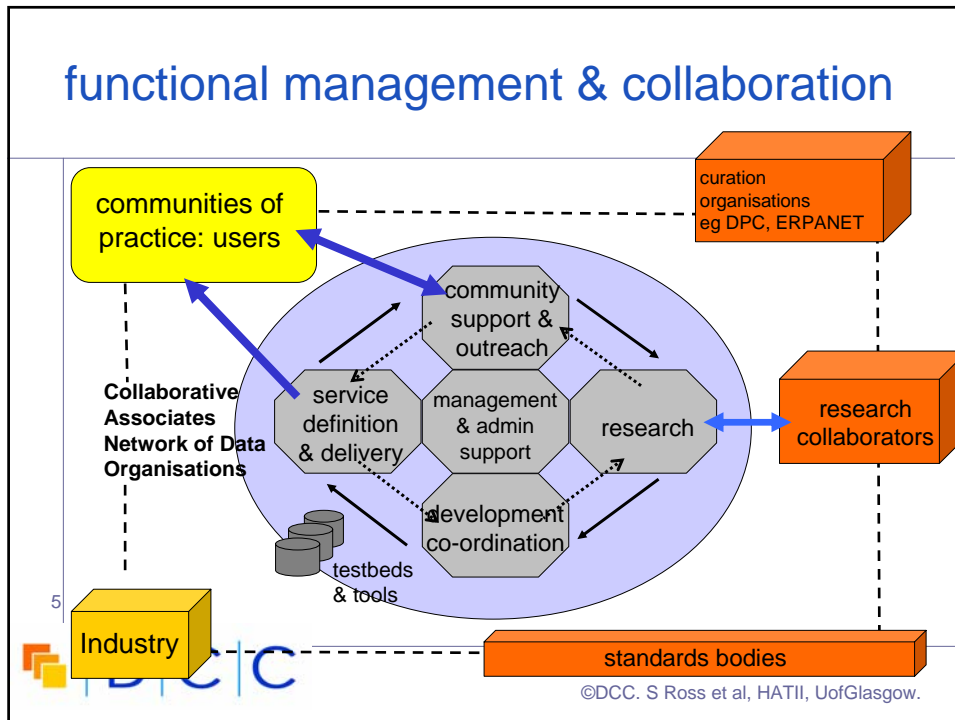
- *The quintessential purpose of the DCC is to support and promote continuing improvement in the quality of data curation and digital preservation activity.*
- *Specifically it..*
 - *Supports preservation of data as an aid to scholarship – reproducibility and reuse*
 - *Enables scholarly communication and eLearning*

3

DCC Objectives

- **Lead a vibrant research programme**
 - addressing the wider issues of digital curation
- **Create an active, innovative, and collaborative Associates Network**
 - strong links across existing community of practice
 - engagement with curators (individuals & organisations)
- **Deliver effective, efficient, and high demand services**
 - to be constructed on cutting edge research, solid developments
 - to include the evaluation of tools, methods, standards and policies
 - to include repositories of tools and technical information
- **Constructing a 'virtuous circle'**
 - built around expertise, experience & requirement feed into the DCC research programme


4



Constructing the DCC

- Initial 3-Year funding stream.
 - JISC began on 1 March 2004
 - EPSRC Research funding began on 1 September 2004
 - expect to harvest 'early crop' from extant research
- Phase One Set-up
 - from 1 March 2004 until launch of Centre in November 2004
 - kick-off meeting, 20-21 March 2004
- For more information, go to Website at www.dcc.ac.uk
 - to learn of updates & progress
 - to access the helpdesk

6



©DCC. S Ross et al, HATII, UofGlasgow.

Data Curation – What?

- actions needed to maintain and utilise digital data & research results over entire life-cycle
 - for current and future generations of users.
- alongside which is Archiving*
 - appraisal & retention/disposal
 - logical & physical integrity: authenticity/security
- and Digital Preservation*
 - long-run technological/legal accessibility & usability
- **Data curation in science**
 - maintenance of body of trusted data
 - to represent current state of knowledge in area of research.

7

Unifying Themes for DCC

- ‘data as evidence’
 - for one or more designated communities
- ‘archival responsibility’
 - at one or more institutional levels
 - with institutional policies & individuals’ competence
- engage/discover communities of practice,
 - invoke/provoke good practices
 - support appraisal & retention/disposal
 - facilitate logical & physical integrity
 - promote authenticity/security
- research problems in productive research domains
 - Informatics, Law

8

Partners

- University of Edinburgh (lead site):
 - Peter Burnhill, Prof Peter Buneman
- University of Glasgow through HATII:
 - Prof Seamus Ross, Director of HATII & ERPANET
- University of Bath through UKOLN:
 - Dr Liz Lyon, Director of UKOLN
- Council for the Central Laboratory of the Research Councils (CCLRC):
 - Dr David Giarretta
- Other key figures include --
 - Malcolm Atkinson, Director, NeSC
 - Chris Rusbridge, Director, Information Services, U of Glasgow

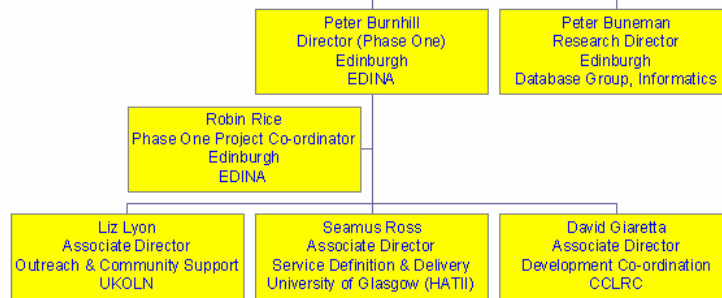
9

Launch Face-to-Face



10

Key Structures



Team manages a rich matrix of activities and personnel.

11

DCC Staffing (ex research)

	Management & Administrative Support	Service Definition & Delivery	Development	Community Support & Outreach	Totals
Yr 1	3.25	5.25	3.5	4	16
Yr 2 & 3	3.25	5.75	3.5	3.5	16

AY 2003/4 £180,141

AY 2004/5 £766,372

AY 2005/6 £779,740

AY 2006/7 £516,242

12


Total for period 1 March 2004 to 30 April 2007 = £2,243,495

Investment in Research

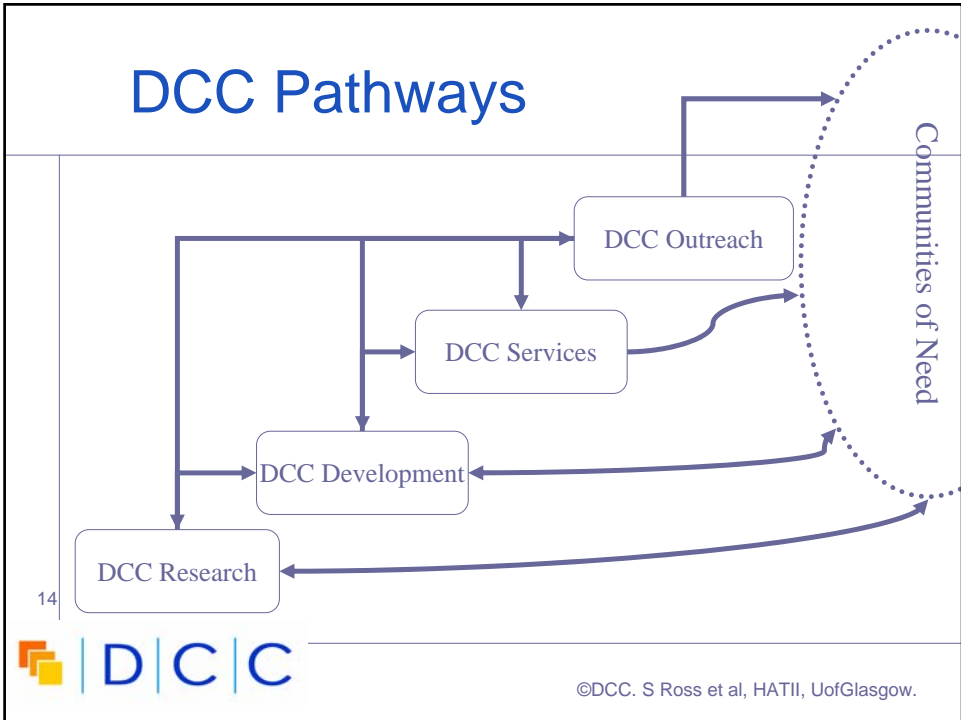
Research funded by EPSRC for period 09/04-08/07

		<i>Staffing:</i>	
Staff:	624,645	UKOLN:	.5
Other costs:	125,000	CCLRC:	.5
Indirect Costs:	187,337	HATII:	1
Total:	1,036,982	Edinburgh:	3
		TOTAL-	6

13



©DCC. S Ross et al, HATII, UofGlasgow.



Research Goals

- Bringing strands of curation, including
 - traditional archival functions
 - maintenance and publication of evolving knowledge as seen in scientific databases
- to identify key challenges and to deliver research results that help address them
- to conduct research in areas crucial to digital curation
- to interlock research activity and service provision
 - Drawn practical challenges to attention of researchers
 - Test research outputs in real world situations

15

Research

- Research
 - Annotation
 - Data integration and publication
 - Appraisal and long-term preservation
 - Socio-economic & legal context
 - rights, responsibilities and viability
 - Performance and Optimisation
 - Semantic metadata extraction

16

Why these research topics?

- they involve fundamental problems
- they will generate highly visible research
- most have components that can be almost immediately carried into development and testing
- of interest to partners and to associates and clients.

17

Development Objectives

- Development planning & monitoring
 - Plan and organise Development thread of the DCC
- Registry/Repository development
 - Define and create registry and repository for metadata
- Representation Information development
 - Development of recommendations for tools and methods for generating Representation Information
- Testing and Certification processes
 - Create a foundation for DCC testing and certification work, including running testbeds

18

Development Deliverables

- **Development planning & monitoring**
 - Produce Development Plan
 - Monitor Development work
 - Seek additional resources for Development work
- **Registry/Repository development**
 - Review registry data model of related projects and agree common model as far as possible
 - Pilot Registry for Data Formats
 - Design of Representation Information Repository
- **Representation Information development**
 - Data Description
 - Software
 - Time Dependence
 - Process
- **Testing and Certification processes**
 - Standards
 - Testbeds
 - Certification

19

Services Objectives

- **Service Planning and Monitoring**
 - Plan and organise Services stream of the DCC
- **Standards, tools, testbeds, and Repositories**
 - Ensure access to standards, tools and repositories of Representation Information
- **Advisory Services**
 - Provide guidance on curation and preservation
- **Audit and Certification**
 - Deliver Audit and Certification Services to the UK HE, FE and research Community

20

Services Deliverables

- **Service Planning and Monitoring**
 - Produce Services Plan
 - Monitor Services Network
- **Standards, tools, testbeds, and Repositories**
 - Develop guidelines on repository development
 - Maintain a standards watching brief
 - Build Catalogue of Tools
- **Advisory Services**
 - Establishing Advisory Software
 - Advisory Services
 - Technical Standards Framework
 - Checklist for Compliance with best practices and standards
 - Current Awareness Bulletin
 - Curation & Preservation Technology & Standards Watch
 - Information, Briefing Papers and Fact Sheets
 - Curation Manual for practitioners
 - An online tutorial
 - Site and Project Visits
 - Regular contribution to discussion lists
- **Audit and Certification**
 - Establish Audit and certification practices
 - Deliver a protocol for audit and certification
 - Development of Certification Procedures
 - Develop self-certification mechanisms

21

Outreach Objectives

- **Outreach planning and monitoring**
 - Plan, manage and monitor the Outreach task area of the DCC
- **Publicity, promotion and dissemination**
 - Raise awareness, promote and disseminate DCC activities.
- **Web portal**
 - Provide support for the community/individuals and provide access to services and resources
- **Training and professional development**
 - Increase the skills and knowledge base of practitioners within different communities of practice.
- **User Requirements Analysis**
 - Gather information about the user / stakeholder community.

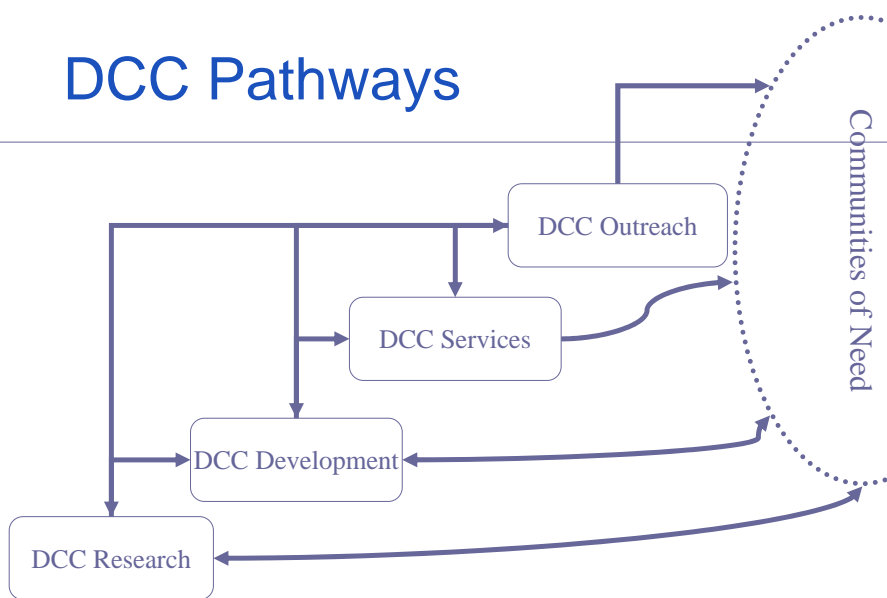
22

Outreach Deliverables

- **Outreach planning and monitoring**
 - Produce Outreach Plan, Manage Outreach activities
 - Review Outreach Plan on an annual basis
 - Bid for additional Outreach funds if appropriate
- **Publicity, promotion and dissemination**
 - Produce publicity materials to include flyer, folder, exhibition stand, posters, standard ppt etc. with appropriate branding
 - Liaise with the media and press offices and provide regular press releases
 - Plan and deliver an events programme to include the formal Launch, an Annual Conference and a series of joint workshops.
 - Provide appropriate communications channels for the community
 - Produce an electronic journal with 4 issues per year to include peer-reviewed material
 - Develop and maintain a contacts database
- **Web portal**
 - Planning and scoping the Web portal project – requirements analysis (Phase 1)
 - Survey/evaluation of available solutions
 - Procurement
 - Design and development including content creation leading to beta implementation
 - Testing, modifications and launch
 - Maintenance, monitoring and extended development (Phase 2)
 - Project management
- **Training and professional development**
 - Deliver a programme of professional development workshops
 - Produce a range of learning materials
- **User Requirements Analysis**
 - Planning, scoping User requirements work
 - Phase 1 User Requirements analysis & gathering

23

DCC Pathways



24

Managing the Initial Challenges of Collaboration

- weekly AccessGrid/telecon for each task area
- Regular face2face meetings
 - Directors
 - Teams
- Forum
- Use of wiki
- Engaging curators in existing community of practice
- Creating Consortium Agreement has proved hard

25

Planning & Progress

- Long-term planning essential
- Substantial territory of action
 - multi-disciplinary, multi data type, multi tradition/profession
 - national and international, but also local and hidden from view
 - Increasing amounts of activity
- Focus on community needs
 - Ensure that we maximise outcome for investment
 - who/what should we plan to affect/effect?
 - policy-makers; 'responsible curators'; (researchers?)
- Collaboration is the key—creating a community of actors
- Official Launch 5 November 2004

26