

## **eBank UK Evaluation Report: Outcomes of eBank UK workshop, University of Bath 5<sup>th</sup> August 2004.**

### **Format and Approach**

The workshop was attended by 25 participants who came from a variety of backgrounds including the Wellcome Trust, the British Library, the Royal Society of Chemistry, the International Union of Crystallographers, CCLRC, the DARE Project, Dspace, and the EPSRC.

Five areas were identified by the Project team before the start as topics where feedback / support were sought. These were:

1. e-Data archives
2. Infrastructure: digital repositories (institutional, disciplinary...)
3. Principle of linking
4. Provide access to your metadata for linking.
5. Mechanism for linking DOI? URL? Ichi code?

The Workshop Programme included three presentations before lunch followed by a plenary feedback and discussion session leading to the Closing Summary. Project staff presented the findings of the project and demonstrated the deposition and aggregator services. The discussion feedback is described below and is listed "verbatim" in note form in three sections:

1. Issues for the JISC
2. Issues for the Project in Phase 1
3. Longer term Issues for the Project

Issues raised will be addressed as far as possible in Phase 2.

### **Issues for the JISC**

A number of issues for the JISC were identified and it was noted that more work is needed in the following areas:

- Developing economic and business models. Additional services mean additional costs.
- Legal / rights issues. Data ownership/IP rights of re-use issues. Investigate use of ?Creative Commons?
- Wider political issues e.g. impact on RAE.
- Government support required following up on House of Commons Select Committee Report
- Developing a national infrastructure.
- Guidance on policies for implementing digital repositories.
- Identification of drivers for successful implementation of digital repositories
- Advocacy to institutions
- Persistent identifiers.
- Sustainability issues and problems with short-term project funding.

### **Issues for the Project in Phase 1**

1. Royal Society of Chemistry "a *tremendous idea*" Noted longevity/sustainability issues. What is the role of learned societies? Global issue.

2. IuCr "*is excited*"; has representation on CODATA  
eBank should consider linking with CODATA
3. Cost benefit issues to the researchers about self-archiving of data. Creates additional bureaucracy/admin when researchers are short of time. Paradox of making self-archiving mandatory with funding vs principle of academic freedom.  
Scalability? Extensibility? RAE issues?
4. Institutional Data Policy + Metadata Policy (public/non-public). Need guidelines on policy for implementing digital repositories.
5. Need national infrastructure of digital repositories.
6. Identify drivers for successful implementation of digital repositories.
7. Digital curation aspects of self-archiving of datasets.
8. Recognise that demonstrators are needed and are valuable to show proof-of-concept and to encourage engagement and buy-in.
9. Advocacy is required by institutions.
10. Develop open tools, open organisations → owned by the community.
11. Netherlands e-data archives  
Issues of Short-term project funding  
Discussion but no infrastructure  
Needs government support  
Report seen as a global driver?
12. Recognise paradox/challenge of Interoperability vs Innovation vs Specific implementation.
13. Disciplinary differences? Multiple links to multiple datasets, multiple formats in social sciences.
14. Social sciences user demand for self-archiving tools.  
There is funder demand (ESRC). Use of national datasets UK Data Archive.
15. Registration issues.
16. Common tools e.g. NESTAR (UK Data Archive).
17. Discovery mechanisms different for Arts + Humanities & social sciences.
18. Data differences: fuzzy/incomplete data  
Errors, validation, data assessment, quality standards.
19. In life sciences, annotations very important, genome data bases, curation issues, provenance issues (starts in lab).
20. Balance between producing workable systems vs perfection?
21. Legislation – H&S requirements relating to labs.
22. Promotion, RAE, impact analysis, careers.
23. Appropriateness of Carrot + stick approach.

## Issues for the Project – longer term

24. Ease of getting data into repository
25. Offer portal of value-added services eg indexing, validation
26. Embedding in CCLRC Data Portal → repository.
27. luCR Perceived barrier = quality of metadata  
DC vs CIF, namespace authority issues, Metadata structures to enable “filtering” (by granularity of topic/discipline). Link with ePrints UK work.
28. Data ownership/IP rights of re-use  
?Creative Commons?
29. Act as a search mirror site?  
Link with luCr
30. Explore wider linking into metadata with institutional .repository @ Southampton e.g. Oceanography.
31. Improve robustness of software to production quality.
32. Provide full text indexing? Googling?
33. Linking CNRI handles  
Persistent identifiers? Risks?
34. Analytical software – custom written. Store programme + dataset. Issues of curation for DCC. Universal Virtual Computer model?
35. Interaction with Dspace
36. In Life sciences annotations are very important, genome data bases, curation issues, provenance issues (starts in lab).
37. Linking results to methodologies, processes + workflow eg synthetic organic chemistry.  
Standards Web Services flow Language  
COSHH forms, Smartlab aims, Ontologies (for lab processes), lchI urls to link (≡RDF).
38. e-Lab environment/position of ISIS in wider science landscape.
39. Provision of user tools to enhance workflow – time issues.
40. M2M generation of metadata & utility.
41. Additional services →COST  
Economic models \$ £ €  
Business models  
Sustainability!