The International Journal of Digital Curation

Volume 8, Issue 2 | 2013

Editorial

Kevin Ashley, **Digital Curation Centre**

This edition of IJDC brings the second and final batch of 15 practice papers from IDCC13, which took place in Amsterdam in January 2013, as well as three research papers submitted directly to the journal. As I remarked in my last editorial, the practice papers are worth your attention, even if you attended that conference, since they allow the authors to explain their work in more detail than is possible in a 15 minute talk.

Eight of these papers describe work carried out under JISC's 'Managing Research Data' programme and other related funding streams in the UK. Although some of the results may be familiar to those who lead in this field, we are aware at IJDC that most research institutions in most parts of the world are still struggling with the early stages of scoping research data management services at an institutional level. We know that making work like this available to a wide audience is thus likely to bring benefits to all those involved.

Garrett, Gramstadt and Silva describe experience with defining requirements and specifying infrastructure for research data in the visual arts – a field in which the term 'data' itself is not widely recognised as being applicable. They include a useful comparison of a number of repository platforms for this particular use case. Even though your circumstances may be different, their approach is likely to be informative. In a related paper, Guy, Donnelly and Molloy describe an approach to defining 'research data' in a meaningful way for creative arts researchers and the institutions they work within. It's important to do this to allow research data services to support the greatest possible number of researchers and to identify those times when it is sensible to provide discipline-specific services. I should note that all three authors were DCC staff at the time that this paper was written.

Four more papers provide insights on different aspects of research data service implementation at university level. Parsons sets out steps taken at the University of Nottingham through the ADMIRe project, including a survey to establish requirements based on the DAF methodology, training for staff and researchers, and the creation of web content to support the services. It ends with reflection on the lessons learned so far in a paper which takes a broad view of work in one university. By contrast, <u>Pink</u> concentrates on one aspect of experience gained at the University of

International Journal of Digital Curation (2013), 8(2), 1-4.

http://dx.doi.org/10.2218/ijdc.v8i2.229

The International Journal of Digital Curation is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/



Bath during the Research360 project. Her description of how Bath worked to reconcile differing funder requirements with 'institutional reality' has relevance far beyond the UK. The paper also contains a useful summary of other results of that project, including the identification of new roles for the emerging research data support service. Wilson and Jeffreys present a follow-up to the University of Oxford experiences described in a paper presented at IDCC10 (Wilson, Martinez-Uribe, Fraser and Jeffreys, 2011). It is thus a useful insight into how plans develop into service roll-out at a world-class research university and contains a good description of various service components and how different parts of the university have worked together to deliver a more joined-up service. In a related paper, Rumsey and Jeffreies describe just one of those components in more detail: DataFinder, the catalogue of Oxford research data outputs. This is an excellent and concise description of the choices made at Oxford about specific issues, such as metadata for research data, and the rationale behind those decisions. It will be useful for anyone embarking on a similar exercise.

Two final papers are not about specific projects in the UK, but instead assess the situation from a national perspective. Molloy et al. describe the approach taken to identify and describe the benefits gained from JISC's funding programmes in this area. It highlights those benefits for which evidence was already apparent, as well as describing future expected benefits for which evidence can only emerge at a later date. That division alone will help others having to undertake similar exercises, whether at institutional, regional or national level. Pryor describes the effects of the closely-related work the DCC carried out over a similar period (and which continues today) providing direct support to institutions to develop research data management capability and translating the lessons learned for others. It highlights the importance of training and also shows how high-level action, such as the Royal Society report 'Science as an Open Enterprise' (2012), can make the process of institutional change easier.

One more paper describes UK work not undertaken as part of the JISC programme. <u>Rice et al.</u> describe another university's move from a research data policy framework to a set of delivered services, this time at the University of Edinburgh where this editor is based. It's a useful framework for others, in particular its recognition that the roadmap for implementation is itself a dynamic plan that is expected to change as a result of its own implementation.

The remainder of the practice papers give us a more global perspective. <u>Schumann</u> and <u>Mauer</u> describe recent work to undertake a formal evaluation of a long-established social science data archive in Germany: GESIS. As well as providing a useful description of the audit process, the paper contains descriptions of the archive's workflow and its relation to the audit process. GESIS is similar in form to many long-established research archives, being a distinct organisational entity embedded in a larger organisation on a single physical site. By contrast, <u>Plale et al.</u> describe a newer approach which is gaining widespread attention: a virtual repository, layered in this case over a number of institutional repositories and potentially using further services, such as cloud storage, to deal with data volumes that the repositories themselves are not designed to handle. The approach taken by SEAD described here has much potential, even if it can only be partially realised. <u>Oostdijk, van dem Heuvel</u> and Treurniet describe the Data Curation Service that operates as part of CLARIN-NL, a national service for research infrastructure in linguistics. As well as a general description of the service, the paper contains an informative worked example of how a particular collection was ingested into the CLARIN-NL infrastructure and made available for reuse.

Halbert presents another institutional picture, this time from the USA at the University of North Texas as part of the DataRes project. As well as describing project findings, the paper looks forward to a future that Halbert is not entirely optimistic about, describing research data management as having a 'problematic future.' Weber et al. tackle one particular challenge of data reuse that many recognise: that of measuring the impact of particular curation activities on data reuse. Their work is based on a study at NCAR, a world-leading large-scale scientific data centre.

In the last of the practice papers from IDCC13, <u>Groenewegen and Treloar</u> set out a national picture of research data management in Australia and describe the work and impact of ANDS, an organisation that we at the DCC view as a valued sibling.

The first of our three research papers continues the theme of research data management at institutional scale and describes the results of research at Emory University into different faculty behaviours and expectations for dealing with research data. <u>Akers and Doty</u> are not the first to carry out such an exercise, as will be clear from their references, but this study is an additional set of data in a field where more research is welcome. The appendix to the paper provides a detailed analysis of the responses to each question.

Bache et al. look in depth at a problem which arises in many areas of digital curation: information provenance. If we create environments which allow information of many types from many sources to be brought together, remixed and repurposed, and then potentially fed back into the same or similar environments, we potentially create significant problems. Information which can no longer be traced to its source can be misleading. In a clinical setting, misleading information can be positively dangerous and it is in the area of clinical research that Bache et al. examine the problem and articulate a model to deal with it.

<u>Moore et al.</u> present a new paper which builds on the iRods system that will be familiar to many readers of this journal. They describe a persistent, distributed archive service based on iRods, which also uses Cheshire3 (a search and retrieval system) and Multivalent (a generic rendering engine).

There's lots to digest there and I know that we already have much more content being made ready for inclusion in IJDC 9(1), including the first material from IDCC14, which takes place in San Francisco in February 2014.

Finally, I should remark on some changes that I heralded in my last editorial for Volume 8(1) of IJDC. We had hoped that those changes would be visible and documented now, particularly those which relate to preserving data behind articles which appear in IJDC. Unfortunately, it hasn't been possible to do this in time for this edition. This year has been a time of significant change for IJDC and many of the staff who do the bulk of the work that allows the journal to appear. A number of readers will be aware of the announcements regarding the withdrawal of the bulk of funding for UKOLN¹; some of the consequences are highlighted in an article from the Times Higher Educational Supplement (Parr, 2013). The journal has also moved homes and hosting environments, from servers at UKOLN to a new hosted journal service based on OJS at the University of Edinburgh. Alex Ball and I gave a talk on the experience and the implications for preservation of journal content at a workshop which preceded this year's Repository Fringe in Edinburgh (Ashley & Ball, 2013). If you didn't notice that anything had changed, then the move went as planned! I'm very grateful to those at UKOLN who made this possible during a difficult period: Emma Tonkin (now with Kings College London), Alex Ball and Bridget Robinson. The OJS team at Edinburgh were also very helpful in making the transition a smooth one. I should also note my ongoing gratitude to Liz Lyon, who frequently steps in this editor's shoes when I am occupied elsewhere, and to Kirsty Pitkin who does an efficient and accurate job with our copy-editing. All of these people do the real work that makes this journal what it is and I'm grateful to them.

References

- Ashley, K. & Ball, A. (2013). Best practice with OJS: A partial view. Paper presented at the Open Journal System Forum. Edinburgh, UK. Retrieved from http://alexball.me.uk/docs/ashley.ball2013bpo/
- Parr, C. (2013). UKOLN hit by job losses as Jisc removes funding. *Times Higher Educational Supplement*. Retrieved from <u>http://www.timeshighereducation.co.uk/news/ukoln-hit-by-job-losses-as-jisc-removes-funding/2003712.article</u>
- Royal Society. (2012). *Science as an open enterprise: Open data for open science* (Royal Society Science Policy Centre report 02/12).
- Wilson, J., Martinez-Uribe, L., Fraser, M.A. & Jeffreys, P. (2011). An institutional approach to developing research data management infrastructure. *International Journal of Digital Curation*, 6(2), 274-287. doi:10.2218/ijdc.v6i2.203

¹ UKOLN: <u>http://www.ukoln.ac.uk/</u>