

# The International Journal of Digital Curation

## Issue 2, Volume 6 | 2011

### Editorial

Kevin Ashley,  
Digital Curation Centre

September 2011

It's pleasing when a problem posed by one paper is apparently answered by another and I'm glad to report that I see some signs of this in the current issue of IJDC. I would be interested to know if the authors concerned agree.

[Botticelli et al.](#) look at digital curation education, and specifically the influence of curriculum choices on the development of the profession. They call attention to a perceived gulf in curation publications between research literature, "often too theoretical" and practice publications that have too narrow a focus. IJDC tries to bridge this gap, but we don't select material with students in mind. However, there's an increasing number of general texts aimed at students who need introductory literature on many aspects of the curation problem. Their call for greater collaboration between research, practice and teaching is one that will find many supporters amongst IJDC's readership. It is one of the functions of the Digital Curation Centre and a principle that is difficult to argue against in many disciplines. So why do we find it so difficult to achieve?

The next three papers meet that need for papers suitable for education as well as for researchers in the field, although it's not the authors' primary purpose.

[Lawrence et al.](#) examine the growing movement for formal publication and peer review of data, and the accompanying need to be able to cite it in a widely-accepted way. In a wide-ranging review of what it means to publish and cite data, and the different ways in which traditional publication can be linked to data citation, they end with a proposal for what they believe is an ideal mechanism for data citation. Although the examples and thinking come primarily from the sciences, much has wider relevance and this paper is recommended reading for those with an interest in the field. It's also an ideal survey of the arguments and the art for students.

[Gregory & Guss](#) provide a useful perspective on education: that of recent graduates from a digital curation course. How do they feel that their course has equipped them for the tasks they must carry out? Are their employers able to utilize the skills they have acquired? I hope that we will see more work addressing this and similar questions in future. In the meantime, read this paper even if you did see their presentation at IDCC '10 in Chicago.

[Higgins](#) provides a historical perspective on the emergence of digital curation as a discipline. Although much will be familiar to regular readers of IJDC, this concise summary of 15 years of work will be a useful reference for many and an excellent introduction for students in the field.

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. ISSN: 1746-8256 The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre.





[Van Horik and Roorda](#) describe the approach taken by DANS to the preservation of databases and similar content. This is essentially conversion-on-ingest (to a long-lived XML format) plus conversion-on-demand at the time of access, with the aim of creating preservation formats which are more long-lived than access formats, whilst preserving significant properties. [Ure et al.](#) describe curation challenges that emerge from another discipline in which research and practice have blurred boundaries - that of telehealth. It is a concern of equal relevance in the developed and developing worlds, and one where patients themselves could make a contribution to the curation of data about them.


[Conway, Giaretta, Lambert & Matthews](#) describe what they term a “preservation” methodology emerging from the CASPAR (and earlier, SCAPE) projects. I quote “preservation” because many will see wider aspects of data management here, which they would term curation. An assertion in the opening paragraphs distinguishes this from much similar work; the authors wish to design AIPs for data which not only make future use more possible, but also more likely. That is, they are seeking to maximize the benefit from the investment in the data’s creation and later management. The paper ends with a recommendation for the DCC and others to develop business cases for the preservation of scientific data. I can assure the authors that the recommendation will be taken to heart. [Khan et al.](#) describe DataStaR’s approach to using semantic web technology and concepts to support data and metadata capture, data sharing, and publication to repositories. Amongst other benefits, this work recognizes the importance of linking institutional research data infrastructure with domain repositories. Aspects of the work have already been taken up outside the project (for example by ANDS), which bodes well for its longer-term viability.

IRods is commonly associated with storage policy, but [Walling & Esteva](#) describe its use to automate the extraction of metadata in digital archaeology. The authors believe that these techniques are applicable outside their domain, and end by making tantalizing reference to possible future work with art photography and herbarium data. [Neuroth, Lohmeier & Smith](#) describe a virtual research environment, TextGrid, which supports the curation of research data in the arts and humanities. TextGrid combines functionality to support active use of data in a lab, and more permanent preservation and access via built-in repository functionality. One hopes that there is a potential for such environments to reduce the costs of many aspects of the curation function. If so, the funding agencies that the authors refer to in their conclusions may be more easily persuaded of the case for long-term support.

Next, a group of papers looking at a variety of institutional perspectives.

[Collie and Witt](#) look at an interesting subset of the institutional data management problem: the curation of data which accompanies doctoral dissertations. Through a case study they provide a compelling argument that this is a good area for institutions to begin to grapple with the problem of managing data for the long term, with material that is more clearly tied to the institution than any that will emerge from other research activities.

[Fear](#) uses a survey of researcher practice in a single institution to consider what services an institution needs to provide to improve practice in research data management. This paper takes the interesting perspective of considering research data as something more-or-less personal – owned, about, relevant to or sent to me, for instance. The conclusion is that treating data management as an activity separate from others is “... confusing to researchers and counterproductive.” We should instead consider it as an aspect of personal information management, with the appropriately broad perspective on what “personal” means in this context.



[Hswe, Furlough, Giarlo & Martin](#) report current practice at Penn State libraries, in particular the creation of institutional services for digital curation and stewardship. The move away from being driven by the capabilities of technology towards a broader consideration of use cases, and how technology might support them, will resonate with many.

[Rice and Haywood](#) outline the University of Edinburgh's approach to dealing with the varying pressures and demands from funders, researchers and others to tackle research data management at institutional scale. The policy which has emerged, summarized in ten concise points, has already drawn wide interest in the UK and elsewhere. Although they have already been referred to elsewhere as "ten commandments" of research data management, this is not a prescriptive policy. Rather, it is a mixture of requirements and promises that are intended to reinforce each other. Work to implement the policy and deliver all aspects of the infrastructure that support it are ongoing, as the authors acknowledge, but the creation and publication of policy is an important step.

[Ward, Freiman, Jones, Molloy & Snow](#) report on the work of the Incremental project, which is acknowledged as an influence by at least one other paper in this issue of IJDC. This project has focused on the creation of non-technical infrastructure to support researchers and its findings have been widely acclaimed even as the project has progressed. Based initially in two institutions, we can already have confidence that the methodology can be adopted elsewhere.

[Wilson, Martinez-Uribe, Fraser & Jeffrey](#)s describe work in a number of projects at Oxford, funded through JISC's Research Data Management programme, which also supported the some or all of the work behind two other papers in this issue of IJDC. The progress they describe is a good example of the progression from project work to institutional embedding, characterized in the five-stage maturity model I referenced in my last editorial. Since this paper was written, a number of these projects have developed into services available outside Oxford itself.

IJDC has previously published a small number of papers on the preservation of digital games, a field which has suffered surprising neglect despite the importance of games as cultural heritage and economic significance. I end with describing one paper in this field, and one which goes beyond it. [Newman](#) looks at the work of the National Videogames Archive with player walkthroughs, but his paper is also a concise summary of the general issues in game preservation. Even more complex is the issue of virtual worlds, which share some of the characteristics of games but bring many more challenges of their own. I am thus very pleased to finally be able to publish the paper by [McDonough and Oldendorf](#), which looks in detail at the many challenges – technical, economic, organizational and legal – involved in archiving the virtual world of Second Life. They are honest about the limitations of what they have achieved so far, but the paper still offers many valuable insights and pointers to future work. We would welcome future submissions on this theme.