Winghardt, Stefan [Editor]; Niedersächsisches Landesamt für Denkmalpflege [Editor]; Institut für Denkmalpflege [Editor]; Puppe, Josefine [Oth.]: Arbeitshefte zur Denkmalpflege in Niedersachsen: Archäologie und Informationssysteme: vom Umgang mit archäologischen Fachdaten in Denkmalpflege und Forschung

Hameln: Niemeyer, Heft 42.2013

Digital curation at the Archeology Data Service

Catherine Hardman, Julian D. Richards

The problems of curating digital archaeological data sometimes tend to be overlooked in larger considerations of an archaeological archives crisis. Compared with the more tangible storage issues facing tons of bulk finds of pottery or the curation of the paper archive, digital data has been regarded as less of a problem, or even as the solution to the problem (Perrin 2002). After all, digital data consumes less space than either finds or paper, and digitisation can provide easier access.

However, it is now recognised that preservation of digital data introduces a whole new set of problems for the curator: data corruption, hardware redundancy, changes in file formats, software migration, and insufficient documentation. It is inadequate to copy data to DVD and put them in the proverbial ‚box on the shelf'. Digital data require active curation. Digital data preservation has developed as a sub-discipline in its own right with its own vocabulary, standards, and organisations. In the UK, the Digital Curation Centre (DCC) and the Digital Preservation Coalition (DPC) are two such bodies that have been working to highlight the digital preservation issues. The call has also been taken up at the European level with a number of collaborative partnerships working together to develop transnational research infrastructures for digital preservation and access and shared standards.

A survey into the digital preservation of research outputs in Europe (Kuipers and van der Hoeven 2009) found considerable awareness of the problem. Preservation is widely accepted to be very important. 91% of researchers surveyed considered the possibility of re-analysis of existing data as the most important driver for preservation, whilst 87% also believe that public funding is a key reason. The same survey, however, found less agreement about who should pay for the costs of digital preservation of research data, with many respondents providing multiple solutions. 61% of researchers believe that national government should pay for preservation, whilst 41% believe that their own organisation should bear the cost.

The problem is particularly acute in archaeology. Archaeological data are increasingly ‚born digital' and represent the primary record of sites destroyed during the development or fieldwork process. Frequently, there are no paper substitutes. Furthermore, archaeologists create large quantities of data and in a greater variety of formats than most other disciplines. (Austin and Mitcham 2007; Austin et al. 2009). Yet, such data are frequently essential to underpin the excavator's interpretation of a site or to allow its reinterpretation by others.

This issue was recognised at a relatively early stage in the UK. The Archaeology Data Service (ADS) was established in 1996 as part of a wider Arts and Humanities Data Service (AHDS). It was charged with preserving research data created within the Higher Education (HE) sector and making them available for re-use for teaching, learning, and research (Richards 2008). Core funding was initially provided by the Joint Information Systems Committee (JISC): the body responsible for UK HE computing and information infrastructure. The Arts and Humanities Research Board (AHRB, later Council – AHRC) later became the primary research funding body for the sector. This meant that University-based researchers could deposit their data with the AHDS which undertook to use its best efforts to preserve these data free of charge into perpetuity.

Core funding continued to support free deposit from the HE sector. As ADS developed, however, it became apparent that it was important for it to develop a business model which could accommodate the preservation of data emanating from the commercial and public sectors, and that derived from archaeological investigations funded by bodies such as English Heritage (EH). A variety of funding models were considered, including charging the users of the data. This was soon dismissed for a variety of reasons. On pragmatic grounds, it was believed that the administrative overheads of charging large numbers of users would not be cost effective. Also, one of the aims of ADS has always been to encourage re-use, not to put obstacles in the way of re-users. Secondly, the data archived rarely have any great commercial value; there was a strongly expressed view that these data had generally been collected using public funding and, therefore, they should be available free to all at the point of use, an approach which was also in line with the broader convention enshrined within the 1992 Valletta Treaty (1) that ‚the polluter pays.'

Therefore, the ADS adopted a business model whereby the costs of archiving outside the academic sector would have to be covered on a project basis. It was also clear that this would have to be a one-off, or one-time only, charge, levied at the point of deposit. Whilst archiving costs would inevitably be ongoing, it was clear that funders would have to budget for the