The idle archive?

written by Jeremy Huggett | 10/02/2017



We often hear of the active archive, but what about an idle one? In a post on Digital Data Realities, I suggested that, although we might wish otherwise, our digital archaeological data repositories seemed relatively little-used. The Archaeology Data Service access statistics did not suggest a large uptake for the project archives it holds, and the ADS had not found it easy to attract entries to its Digital Data Reuse Awards in the past. In that light, I commented that it would be interesting to see how the OpenContext & Carleton Prize for Archaeological Visualization would get on. Well, the jury is now in, and the winner is … the 'Poggio Civitate VR Data Viewer', an impressive-looking data viewer, though as it requires an HTC Vive to use, I can sadly only watch the video rather than experience it myself …

However, as interesting are Shawn Graham's reflections on the experience of organising the contest:

> "We offered real money – up to a $1000 in prizes. We promoted the hang out of it. We made films, we wrote tutorials, we contacted professors across the anglosphere. We had very little uptake."

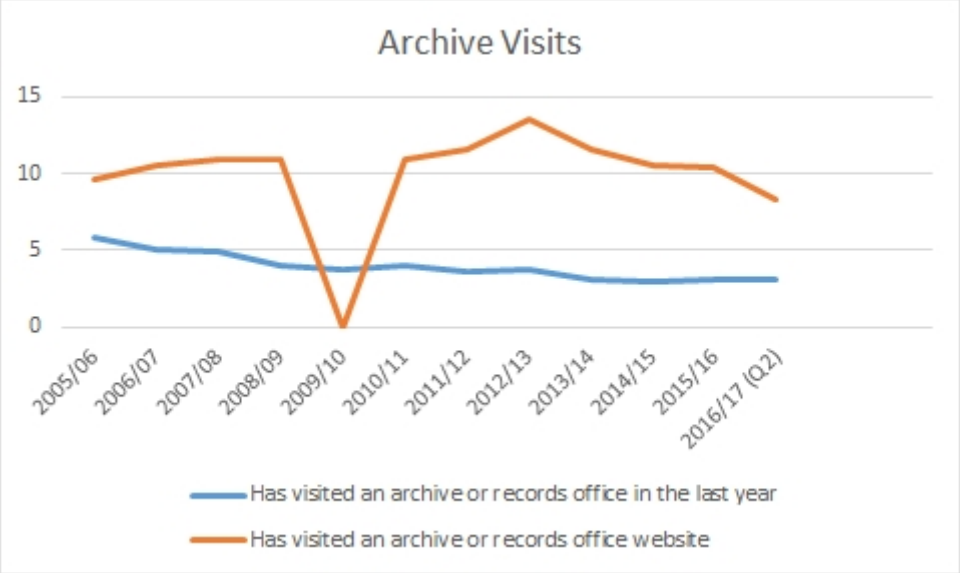
(accompanied in his presentation by an image of tumbleweed) … Indeed, only the one winner was announced for the team prize – no individual or student prizes were awarded as was originally intended. So what's going on?

We presumably accept the importance of long-term data retention and the value of archaeological digital archives, we accept the ethical obligations to archive our primary data, and considerable financial and personal resources are expended to these ends, yet the uptake of the resulting datasets seems limited at best. We put the data in, but no one appears to want to take it out.

To some extent, we shouldn't perhaps be surprised. As far back as 1999, Nick Merriman and Hedley Swain pointed to the lack of use of physical archaeological archives. They showed that 29% of museums surveyed received no visits or inquiries to their archives in the previous year, and that annual total visits to archives averaged 46 (1999, 260). They described archives as intractable and inaccessible, and pointed to the potential of digital technology for opening up such archives to greater use. In some respects, not a lot seemed to have changed by the time of the Society of Museum Archaeologists' *Archaeological Archives and Museums 2012* report which had the greater promotion of the potential of archaeological archives as a resource for engaging communities as one of its key recommendations (Edwards 2012, 9). In a subsequent open meeting to discuss the

recommendations presented in the report, the lack of visibility of resources was frequently raised, and the importance of public access to the relevant software (and presumably the knowledge to use it) was also noted.

Nor is this a particularly archaeological problem. A research study published by The National Archives, *Digital services and archive audiences* (White, 2014), identified a decline in visits to local authority archives between 2005 and 2014 which seemed partly offset by a rise in users of archive websites. The conclusion drawn was that existing users of archives were turning to online services for access and information (Thomas 2014, 9). On the other hand, the Taking Part survey, the main evidence source for the Department of Culture, Media and Sport in England, suggests a rather different picture: although more people visit archive websites than the physical archives themselves, the trend over the same period is flat if not downward even after allowing for a peak in 2012/13 thought to be linked to the London Olympics (see graph below). So it would seem that we cannot necessarily rely on a digital 'bounce' as users turn to digital archives instead of their physical counterparts.



Archive Visits in England (percentage data derived from the Taking Part Survey)

Of course, one of the problems with these kinds of data is that the idea of a user 'visit' is pretty meaningless in digital terms. While it may be a reasonable assumption that a visit to a physical archive is accompanied by the use of that archive to some degree, a visit to an archive's website may be no more than that, and certainly cannot imply the download and subsequent use of a dataset. Consequently, the actual use of the digital archive may be considerably less than visitor numbers might suggest – something that chimes very much with the archaeological digital archive experience.

A low level of use of archaeological digital archives might not therefore be out of line with the broader picture elsewhere, but that is small comfort. So what might be going on? Shawn Graham suggests one of the primary reasons why archaeological digital data are not being rehashed is a lack of appropriate levels of digital literacy. He sees this in terms of a lack of a culture of

undergraduate teaching using actual datasets and an associated fear of failure amongst the students if the digital tools/data don't work or don't generate the expected outcomes. The need for a degree of digital literacy was also highlighted in the discussions surrounding the SMA report on archaeological archives. Although I teach a course on GIS in archaeology using datasets that students acquire from the ADS and elsewhere, it is an optional course taken by at most 25% of eligible students and, apart from one or two dissertation projects, the remainder are unlikely to get equivalent exposure to the acquisition and processing of archived data elsewhere. So there is a literacy gap. However, where the appropriate level of literacy sits is a rather different matter. For instance, the OpenContext & Carlton Prize had a set of requirements significantly more advanced than many students would likely have, or most archaeologists or members of the public. These included the ability to mount visualizations on the open Web, to create and/or incorporate source code and make it available via Github etc., to utilise the OpenContext API rather than simply download data for subsequent reuse, and so on. A much lesser level of literacy is necessary for standard access to the archives, thank goodness, but it perhaps suggests that the real problems likely lie elsewhere.

A key question is what are we using these data for (or, since we're apparently not using them, what should/could we be using them for)? Teaching is certainly one area, although by no means all students will experience this, as we've seen. Experience with physical archives suggests that a high proportion of their usage is associated with displays and exhibitions, although the next largest categories were by researchers and for education (largely school) purposes. Archive resources were rarely used by commercial archaeologists and overall the actual instances of use were numerically low (Merriman and Swain 1999, 259-60). These conclusions were broadly confirmed in the SMA report (Edwards 2012, table 7). By these measures, digital archives would seem to have relatively low potential for displays and exhibitions, but what about the other categories of use? We might expect greater use by commercial sector, but arguably secondary sources and the higher level catalogues rather than archive data would be more likely consulted in desk based assessments and the like. That leaves researchers …

So why are researchers not making more use of digital archive data? It may simply be that our expectations of archive use are unrealistic – a kind of 'if we build it they will come' approach, whereas the reality is that they will come, but in a generation or so, when it comes time to re-evaluate or reappraise a project or programme of work or, indeed, a researcher. So in other words, we may be worrying too soon. On the other hand, there may be an issue with perceptions surrounding the creation of new data versus the recycling of old data. It has proved challenging within current university environments to receive credit for the creation and archiving of data in the first place, despite the efforts of the Journal of Open Archaeology Data or Internet Archaeology in developing 'data papers', for example, or the recognition in other disciplines of the value of the products of research alongside more traditional academic outputs. At the same time, it may be perceived that it is somehow more important to generate our own new data and draw our new conclusions from them than it is to rework old data in search of those new conclusions. This may be a consequence of human nature, peer pressure, research review criteria, or whatever, but it comes through in Ethan Wattrall and Shawn Graham's topic model of archaeological journal articles which from the terminology used seems to illustrate a reluctance to analyse other people's data. However, one aspect of data recycling which might sidestep such perceptions is where different datasets are combined to generate new information, but here the issue might be a relative lack of data capable as yet of being meaningfully mashed together. Nor is this a trivial process – bringing together data

collected at different times, by different people, for different purposes, and under different recording regimes, is challenging at best (for example, see Allison 2008).

Should digital archives really be any different to their physical counterparts in terms of their usage? We tend to assume that the greater accessibility and flexibility offered by the digital tools should in and of themselves attract users, and are puzzled when that is not evidenced to be the case. Of course, in an ideal world, the relative lack of use of these archives is not a great problem – they continue to serve the important purpose of ensuring that valued and valuable data are preserved long-term, and that fragile digital information is maintained and migrated through time such that they retain their potential for use at some point in the future. However, in the real world the value of these archives is most easily demonstrated through their use now, and hence provides vital justification for the resources sought for and committed to their maintenance. So we are left with a conundrum.

Strange as it might seem, it may be that in addressing this question of the use of digital archaeology archives, we are too focussed on the archive as a repository of data. This is inevitable, as we have been engrossed in the phase of creation and construction of these archives, on issues of standards, migration, and the mechanics of search and retrieval. Only now, as they become increasingly established and populated with datasets, are issues concerning their use (or lack of it) coming to the fore. We may instead need to stand back and consider the archive in the broader sense, as a subject in itself. For example, Baird and McFadyen argue that a shift in perception is required, from using archives as a primary source to studying them as a topic in themselves (2014, 14). Although their emphasis is on the creation of archives during the excavation and post-excavation processes, here the suggestion is that we refocus on the creation processes behind the construction of our digital archives themselves. As Baird and McFadyen suggest:

> "The form of the archaeological archive, we believe, is related directly to the form of archaeological knowledge. The recognition of this relationship is key to working with 'legacy' data …" (2014, 15).

As digital archaeological archive data are by definition 'legacy' in some sense, considering the relationship between these data and archaeological knowledge creation may help shine a light on their role and place in archaeological practice and provide reassurance about their value now and in the future.

References

P. Allison 2008 'Dealing with Legacy Data – an introduction', *Internet Archaeology* **24**. http://dx.doi.org/10.11141/ia.24.8

J. Baird and L. McFadyen 2014 'Towards an Archaeology of Archaeological Archives', *Archaeological Review from Cambridge* 29 (2), 14-32. https://www.researchgate.net/profile/J_Baird2/publication/291943480_Towards_an_archaeology_of_archaeological_archives/links/573ee1b408ae298602e8e0ce.pdf

R. Edwards 2011 *Archaeological Archives and Museums 2012* (Society for Museum

Archaeologists). http://socmusarch.org.uk/socmusarch/gailmark/wordpress/wp-content/uploads/2016/07/Archaeological-archives-and-museums-2012.pdf

S. Graham 2017 ‘Rehashing Archaeology – Critical Perspectives in Digital Archaeology’, *Electric Archaeology*, Feb 3 2017. https://electricarchaeology.ca/2017/02/03/rehashing-archaeology-critical-perspectives-in-digital-archaeology-feb-3rd-2017/

N. Merriman and H. Swain 1999 ‘Archaeological Archives: Serving the Public Interest?’, *European Journal of Archaeology* 2 (2), 246-267. http://dx.doi.org/10.1179/eja.1999.2.2.249

T. White 2014 *Digital services and archive audiences: Local authority archives 2014* (The National Archives, Richmond, Surrey). https://www.nationalarchives.gov.uk/documents/archives/Digital_Services_and_Archive_Audiences_2014.pdf