

# Archiving London's digital archaeological data: problems, pragmatism ... and progress?

Edward Evans

## Introduction

Throughout 2005/6 LAARC (the London Archaeological Archive and Research Centre) has reworked its standard for the deposition of archaeological archives;<sup>1</sup> a set of criteria that must be met for the acceptance of material into the collection. The previous edition of these standards was issued in 1998. Since this time information and communication technology (ICT) has transformed the information archaeologists create, the way they create it, how they keep it and how they publish it.<sup>2</sup> The reworking of the *Standard* may have its genesis in technological change, but it has implications that extend beyond the digital; encompassing paper and photographic records, and the strategies employed in collecting, storing and disseminating knowledge about London's past.

Through this short paper I use the experience at LAARC and the timely release the Archaeological Archives Forum's 'Review'<sup>3</sup> to explore the tensions inherent in the creation and negotiation of workable deposition standards; and, more significantly, how this potentially informs the histories that archivists present to their diverse audiences. The intention is not to critique any particular standard, but to solicit feedback for a subsequent publication.<sup>4</sup>

## The problem

In 2000 the UK Government<sup>5</sup> anticipated that "the archive sector will adapt itself to the information revolution". At the same time, Richards<sup>6</sup> succinctly characterised the changing mechanisms through which archaeologists were committing their finite and diminishing resource to these same custodians;<sup>7</sup> commenting, "modern archaeological fieldwork projects create huge amounts of digital information ... it is possible to

create digital information at every stage from initial desk-based assessment to publication". By contrast, only two years previously Swain's<sup>8</sup> survey of archaeological archives suggested that most had encountered little difficulty dealing with digital material, primarily because few held digital data.

The effective management, preservation and dissemination of digital data has, however, become a significant concern for many archives.<sup>9</sup> Most simply, archives have been compelled to respond to the transformations identified by Richards. The impetus within the archive sector, however, may be more subtle, but nonetheless is equally persuasive and pervasive.

Swain<sup>10</sup> observed that "archaeological archives are under-used and are not being exploited to their full potential".<sup>11</sup> This is in stark contradistinction to the *raison d'être* of many archives such as LAARC whose mission statement, *Making London's Archaeology a Living Resource for All*, sits alongside performance targets that seek a five-fold increase in users between 2004 and 2009. ICT, particularly the World Wide Web, has been readily seized upon by many as the panacea for resource dissemination<sup>12</sup> and has ultimately become enshrined in government policy.<sup>13</sup>

The powerful and populist appeal of networked delivery has created an environment in which technical, legislative and social imperatives have culminated in an archive/ICT interface that is user focused and, arguably, privileges the front-end of archival practice. The former is clearly laudable, but the latter requires more critical appraisal. Simply, although it is widely recognised that "the core of archaeology lies in the long-term preservation of the past for present and future benefit",<sup>14</sup> the shift to digital records entails entrusting more and more primary data to a form

that “is notoriously fragile, and ... tends to quickly deteriorate and become obsolete”.<sup>15</sup> If this fragile data is not made secure, what will archives disseminate? The back-end of practice requires as much attention as the front.

## **A pragmatic approach to digital archiving**

A bewildering set of initiatives have characterised the role of ICT within the cultural heritage sector over the last 20 years,<sup>16</sup> with records created in a plethora of types, formats and structures;<sup>17</sup> a problem exacerbated by the diverse nature of archaeological fieldwork. In contradistinction to this diversity, archaeological practice requires that all “projects must result in a stable, ordered, accessible archive”.<sup>18</sup> Data standards are one means of allowing the receiving archive to establish control over the material it is to manage<sup>19</sup> and bring the stability demanded.

In the most recent appraisal of archaeological deposition standards, Brown,<sup>20</sup> contra Keene,<sup>21</sup> contends that digital data remains a “severe problem area”. Ross<sup>22</sup> has suggested that “apathy, lack of realisation of the urgency of the digital data problem, and a shortage of skills is leading to inaction”. Certainly it is not unreasonable to argue that many archaeological repositories still do not have the appropriate technology or training to store, access and curate digital material,<sup>23</sup> but papers presented at the recent *Society of Museum Archaeologists* conference,<sup>24</sup> strongly suggest that the problem is neither apathy nor ignorance but implementation; implementation that requires sufficient resourcing.<sup>25</sup> Where funding is secured it is often piecemeal, derived from a succession of individual awards, or targeted at eye-catching inclusion and access initiatives.<sup>26</sup> This is a long way from the picture of sustainable, trustworthy repositories promoted by the digital archiving community itself.<sup>27</sup> In an environment where extensive in-house data processing is precluded by limited resources and in which new deposits are constantly being generated, it is vital for the sustainability of the resource that data is accessioned in appropriate formats and structures, with sufficient documentation and metadata to ensure its future reuse.

A repository may issue standards and guidance, but just as an archive requires adequate resources,

so “standards can not be met without resources for the job”;<sup>28</sup> whether this is for the resource-hungry creation of metadata and documentation, or adjustments to established workflow trajectories.<sup>29</sup> Consequently, in many ways archives need the consent of their depositors to fulfil deposition requirements; depositors who may embrace the “highest of the high tech”,<sup>30</sup> lack significant technological expertise, or work under more than standards regime. Archive users, in turn, require assurances of quality, integrity and authenticity.<sup>31</sup> Consequently, in developing new digital standards, LAARC has embarked upon an extensive consultation exercise. The result is a pragmatic standard that seeks to balance the different skills and needs of the archive, its depositors and its users.

## **Does pragmatism meet our archiving needs?**

It is reasonable to suggest that taking a pragmatic approach to standards development is likely to produce a workable standard that facilitates the accessioning of archives. Just as it is true that “all aspects of the archaeological process affect the quality of the resulting archive”,<sup>32</sup> then, logically, taking pragmatic decisions at the point of ingest must hold implications for the subsequent management and dissemination of the resource. For brevity, the following discussion is confined to data *per se*, but the arguments espoused should be understood in the context of significant metadata and documentation requirements.

Pragmatism may be manifest in many different guises within a deposition standard. On occasion, it may, for example, be advantageous to be more insistent on structure than on format. That is, appropriately structured data may facilitate future migrations with greater ease, and therefore less expense, than retrospectively re-working poorly constructed data regardless of format. Given the published specification and ubiquity of use, LAARC has taken this approach to the archiving of GIS data as ESRI shapefiles. Not all formats, however, can be attributed the same short-term confidence as the shapefile, and certainly pragmatic format choices should not be regarded as engendering the same long-term security inherent in recognised preservation formats. Pragmatism, in this context, may consequently be

viewed as simple deferment, but it is also a strategy that holds more ephemeral, and wide-reaching implications; implications that can be most succinctly addressed through consideration of digital photography.

In recent years, digital photography has become a somewhat contentious issue in archaeology. Concomitant to illustrating the attractions of digital photography to the archaeologist, Wooliscraft<sup>33</sup> has effectively critiqued the technological limitations of this method of data capture, limitations that also offer fundamental challenges to the archivist.

The archiving of digital images adheres to the central precept of capturing high quality, stable master images from which deliverable surrogates are derived.<sup>34</sup> The capacity to capture potential archival masters currently remains a feature of more high-end, and hence expensive, cameras with attendant repair and replacement costs. [Similarly, strategies that employ some form of in the field data processing carry their own additional expense]. Understandably, contractor preference is for the continued use of lower cost cameras that generally employ JPEG as the standard storage format.

One pragmatic solution might be to offer a sliding scale of preferred deposition formats with the archive subsequently batch-converting non-archival formats into preservation versions. JPEGs, however, utilise a form of lossy compression (the exclusion of data to reduce file size), that cannot be recovered by conversion to more stable formats – a particular concern for archaeologists given Wooliscraft’s emphasis of appropriate resolution and the subtleties of tonal change. Contra the maxim “ICT must serve not control”,<sup>35</sup> in this context it does not seem unreasonable to suggest that the accessioning of JPEGs constitutes a clear case of technology driving practice; of short-term economisation overriding the expense of future problems and potentially prejudicing the longevity of the data.<sup>36</sup> Consequently Brown,<sup>37</sup> echoing Wooliscraft, comments “digital photographs are not part of the documentary archive ... They do not have proven archival longevity and must not be used as a substitute for black and white record photography”. Given this assertion, the convenience of digital photograph for the

contractor could be pragmatically balanced with the archive’s need for longevity in a standard that mandates the submission of black and white negatives for the ‘last will and testament’ of significant features and allow the deposition of JPEGs for other site shots.<sup>38</sup>

By contrast, at the other end of the archival process users are increasingly expecting information to be available online,<sup>39</sup> and that this information will have been enhanced, improved and therefore of better quality than material not accessible in this manner;<sup>40</sup> an impression enhanced by initiatives that encourage archives to digitise their most popular material.<sup>41</sup> Extending the digital photography analogy, for example, and operating in an economic environment that precludes expensive day-to-day digitisation programmes, it is the ‘incidental’ site shots, rather than those of recognised archaeological value that have pre-eminence in the eyes of the network-focussed user. In this context, in ensuring that the fragile social memory of archaeology is not lost, pragmatism may constitute little more than a dilution of the information and educational potential of the archive by misrepresenting the significance of its components within established understandings of value. Describing the use of digital photography as such is overly dramatic, but this example demonstrates that in a very real way taking a pragmatic approach to the creation of deposition standards determines what enters the archive in digital form and in turn the perceived value of the archive.

## Progress?

Taking a pragmatic approach to deposition standards can be argued to create a rather backward-looking standard that potentially prejudices the perceived value of an archive. This is not to argue that there is not a place for pragmatism, nor to claim that being resistant to the introduction of technology is little more than standing in the way of progress, but rather to suggest that in a situation where in-house data enhancement is precluded, pragmatism remains necessary but needs to be implemented with a mind as much to the future as the past. Institutions such as the ADS are acquiring a history that demonstrates that electronic archives are as sustainable as their hardcopy or film

(including microfilm)<sup>42</sup> counterparts and this should be recognised in the creation of any archaeological standard.

There can be little doubt that our society will continue to witness rapid changes in technology, taste and markets.<sup>43</sup> But, as the cost of what is currently high-end technology continues to fall and, somewhat paradoxically, technological innovations such as XML-enabled office products and MIDASXML, for example, become more widely used, such change can hopefully be delivered in a cost effective, stable and sustainable manner; facilitating the development of digital archives (at a national, regional or local level) that can be forward- rather than backward-

looking. Perhaps the time is coming when archaeological archives will be able to move away from the black and white negative and more fully embrace the digitally-born era.

The question remains as to what role archaeological archives should play in facilitating this fundamental shift; to ask at what point should deposit standards move to reflect, or maybe even encourage, this sea-change? If the profession is to be primarily user-driven, are archives compelled to argue for an enhanced role for digital data capture, storage and transmission? But if so, is it possible to have confidence in the facilities and skill-sets currently available to effectively curate such material?

1. Expected publication date autumn-winter 2006.
2. See D. Carlston 'Storing Knowledge' in M. MacLean & B. Davis (eds) *Time and Bits: Managing Digital Continuity* (1998) 21–31; S. Ross *Changing Trains at Wigan: Digital Preservation and the Future of Scholarship* (2000) 2.
3. D. Brown *Recommended Standards for the Creation, Preparation, Deposition and Curation of Archaeological Archives in England, Northern Ireland and Scotland*. Consultation Document available from the Archaeological Archives Forum <<http://www.britarch.ac.uk/archives/>> (2006).
4. E. Evans, in prep. *A Pragmatic Approach to Archiving Archaeological Digital Data*.
5. UK Government 'Government Policy on Archives' *J Soc Archivists* 21 (1) (2000) 11–23.
6. J. Richards et al. *Digital Archives from Excavation and Fieldwork: Guide to Good Practice* (2000) accessed July 2006 <<http://ahds.ac.uk/creating/guides/index.htm>>.
7. See also B. Davis and M. MacLean. 'Mapping the Project, Grasping the Consequence' in M. MacLean and B. Davis *op cit* fn 2, 1–10; H. Swain *A Survey of Archaeological Archives in England: A Report for English Heritage and the MGC* (1998); Ross *op cit* fn 2, 2–3; UK Government *op cit* fn 5, 13; PRO. *e-government Policy Framework for Electronic Records Management* (2001) accessed July 2006 <[www.e-envoy.gov.uk](http://www.e-envoy.gov.uk)>.
8. Swain *op cit* fn 7.
9. See C. Pickford *Eastern Promise: A Strategy for Archival Development in the East of England* EERAC (2002); K. Barata 'Archives in the Digital Age' *J Soc Archivists* 25 (1) (2004) 63–75; TNA. *Report on Responses to The National Archives Consultation Paper CP03/01: Proposed National Records and Archives Legislation: Proposals to Change the Current Legislation Provision for Records Management and Archives* (2004) accessed July 2006 <<http://www.nationalarchives.gov.uk/policy/proposed/pdf/report.pdf>>; Brown *op cit* fn 3, p.1.1.2.
10. Swain *op cit* fn 7, 9.
11. See also Irvine 'Government Policy on Archives.' *J Soc Archivists* 21 (1) (2000) 11–12.
12. See S. Keene *Digital Collections: Museums and the Information Age* (1998) 9; Swain *op cit* fn 7; A. Kenney and O. Rieger 'Introduction: Moving Theory into Practice' in A. Kenney and O. Rieger (eds) *Moving Theory into Practice: Digital Imaging for Libraries and Archives* (2000) 1–10; UK Government *op cit* fn 5; NCA. *Archives in the Regions: An Overview of the English Regional Archive Strategy* (2001); re:source. *Information and Communications Technology and the Development of Museums, Archives and Libraries: A Strategic Plan (Draft)* (2001); Pickford *op cit* fn 9; M. Moss and J. Currall 'Digitisation: Taking Stock' *J Soc Archivists* 25 (2) (2004) 123–37; H. Forde 'Access and Preservation in the Twenty First Century: What has changed?' *J Soc Archivists* 26 (2) (2005) 193–200; D. Griffin and Halpin 'An Exploratory Evaluation of UK Local e-government from an Accountability Perspective' *The Electronic Journal of e-government* 3 (1) (2005) 13–28, accessed July 2006 <[www.ejeg.com](http://www.ejeg.com)>; NSB. *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century* (2005) 4–5, accessed 2006 <[http://www.nsf.gov/nsb/meetings/2005/LLDDC\\_draftreport.pdf](http://www.nsf.gov/nsb/meetings/2005/LLDDC_draftreport.pdf)>.

13. Initiatives such as the *Modernising Government* agenda and the *Freedom of Information Act* typify an approach to government that is intended to be citizen orientated, transparent and available 24/7 (see NCA *op cit* fn 12; PRO *op cit* fn 7; Barata *op cit* fn 9; A. Allan *Invited paper: e-government information: the same old problem ... newly digitised!* Social Sciences Online (2005) accessed July 2006 <<http://www.sosig.ac.uk/socsciweek/blog/?p=24>>; L. Damodaran et al 'The Contribution of Sociotechnical Systems Thinking to the Effective Adoption of e-government and the Enhancement of Democracy' *The Electronic Journal of e-government* 3 (1) (2005) 1–12, accessed July 2006, <[www.ejeg.com](http://www.ejeg.com)>.
14. C. Perrin *Archaeological Archives: Documentation, Access and Deposition, a way forward* AAF (2002) accessed July 2006, <<http://www.britarch.ac.uk/archives/>>.
15. B. Lavoie *The Incentives to Preserve Digital Materials. Roles, Scenarios and Economic Decision Making*. OCLC (2003) accessed July 2006 <<http://www.oclc.org/research/projects/digipres/incentives-dp.pdf>>; P. Lyman and H. Besser 'Defining the Problem of Our Vanishing Memory: Background, Current Status, Models for Resolution' in M. MacLean and B. Davis *op cit* fn 2, 11–20; Richards et al *op cit* fn 6, 1; Ross *op cit* fn 2, 2; PRO *op cit* fn 7, 7.
16. J. Hemsley et al 'Conclusions and Future Trends' in J. Hemsley, V. Cappellini and G. Stanke (eds) *Digital Applications for Culture and Heritage Institutions* (2005) 295–300.
17. Ross *op cit* fn 2, 9–22.
18. Brown *op cit* fn 3, 1.2.5.
19. S. Granger et al. *Cost Elements of Digital Preservation*. University of Leeds (2000) accessed July 2006 <<http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc>>; A. Brown *Digital Archiving Strategy* English Heritage (2002) accessed July 2006, <<http://www.english-heritage.org.uk/server/show/nav.001002003009006002>>.
20. Brown *op cit* fn 3, 1.1.2.
21. Keene *op cit* fn 12, 17.
22. Ross *op cit* fn 2, 19; also Y. Lusemet and P. Drenth 'Preservation and Access: Two Concepts, one goal - The work of the European Commission on Preservation and Access (ECPA)' *J Soc Archivists* 20 (2) (1999) 161–8.
23. Swain *op cit* fn 7, 47; PRO *op cit* fn 7, 7; Pickford *op cit* fn 9.
24. Museum of London June 2006.
25. See also NCA *op cit* fn 12, 5; TNA *op cit* fn 9, 19.
26. Moss and Currall *op cit* fn 12, 126.
27. See RLG. *Attributes of a Trusted Digital repository* RLG-OCLC (2002) accessed July 2006, <<http://www.rlg.org/longterm/attributes01.pdf>>; B. Lavoie *Technology Watch Report. The Open Archival Information System Model: Introductory Guide* DPC (2004) accessed July 2006, <<http://www.dpconline.org/docs/reports/dpctw05-01.pdf>>; A. Brown *Cost Modelling: the TNA Experience. Paper presented at the DCC and DPC Workshop on Cost Models for Preserving Digital Assets, 26th July 2005*. The British Library: London; R. Heery and S. Anderson. *Digital Repositories Review*. JISC (2005) accessed July 2006 <[http://www.jisc.ac.uk/uploaded\\_documents/digital-repositories-review-2005.pdf](http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf)>; NSB *op cit* fn 12.
28. Pickford *op cit* fn 9, 4.
29. See, for example, Granger et al *op cit* fn 19, 10; H. James *Guidelines for Documenting data AHDS* (2004) accessed July 2006, <<http://ahds.ac.uk/creating/information-papers/index.htm>>.
30. D. Lees 'Lasers and the Lord Mayor's Coach' *Current Archaeol* 202 (2006) 548–52.
31. Moss and Currall *op cit* fn 12, 136.
32. Brown *op cit* fn 3, 1.5.1.
33. D. Wooliscraft 'Thoughts on the Suitability of Digital Photography for Archaeological records' IFA (2005) accessed July 2006, <<http://www.archaeologists.net/modules/icontent/inPages/docs/DW%20digifilm.pdf>>.
34. Richards et al *op cit* fn 6, 4.3; Moss and Currall *op cit* fn 12, 126.
35. re:source *op cit* fn 12, 8.
36. See Lyman and Besser *op cit* fn 15, 13; Ross *op cit* fn 2, 6.
37. Brown *op cit* fn 3, 3.2.6.1.
38. See also Wooliscraft *op cit* fn 33, 6.
39. Kenney and Rieger *op cit* fn 12, 1. See also discussion of project funding above.
40. Allan *op cit* fn 13, 1.
41. E.g., UK Government *op cit* fn 5, 19.
42. See Moss & Currall *op cit* fn 12 for discussion.
43. Hemsley et al *op cit* fn 16, 297.