4   The ADS: What Is It All About?

Stuart Jeffrey (`mailto:stuart.jeffrey@york.ac.uk`)

Everyone in York knows about the Archaeology Data Service (ADS, `http://tinyurl.com/78p4d6f`). We are the nice people in the offices near the entrance to the King's Manor. Hold on a minute though, what exactly is the ADS? What do we do? To start to answer these questions it is necessary to take a little step back and ask ourselves a very important question. What is the purpose of archaeology?

Clearly there is neither a simple, correct, nor even single answer to this question, but there are lots of good and interesting answers that range from the very philosophical to the very practical. Many (not all) of these answers do have one element in common though. This is the notion that archaeologists create as well as interpret information and the information that archaeologists create, the data they gather, should be made available to others for reuse in the future. This can be said to apply to many disciplines, but it is especially important for archaeology because, as we know, some archaeological processes, such as the process of excavating a site, are destructive and unrepeatable acts. What remains after the site is excavated are the memories of the experience, the impressions of those affected by the site and the ideas about the past that those involved in the work – and those watching it happen – have created through direct contact with and through consideration of the material that has been uncovered. Once the project is complete the main connection back to the site apart from memories and selected physical remains are the records that are generated throughout the archaeological process, the primary data, and the ideas about people in the past that these records have helped to inform, the interpretation.

There is much focus on the publication of a site or project, and rightly so for two reasons. Firstly, it is important that archaeologists share their ideas with colleagues and the public, and secondly, academics are often judged on their publication record. For aspiring academics it is crucial to their career that they publish. However, what becomes of the bulk of the material on which their publications are actually based, the material that they often destroyed a site to create? Clearly it is important for archaeologists that these records are kept safe for the long term. This is often an explicitly stated aim of archaeological projects under planning legislation and also as a fundamental part of professional good practice. All the carefully constructed records, the geophysics plots, the context sheets, the photographs, are intended to be kept safe for the long term so that others can revisit them in the light of new discoveries, or developments in technology or theory, in order to challenge the original interpretation and to generate new ones. It is also important to understand that this is not a simple bureaucratic or administrative task. Our entire understanding of the way perceptions of the past have evolved from the earliest antiquarians to the modern discipline are contingent on both the published material and the unpublished archival material that can complement, or contradict it. Understanding how the archive is created, the selection and retention processes involved, which are themselves an act of interpretation, are an intrinsic part of being an archaeologist.

So how do archaeologists fulfill their responsibility to keep their records safe for the future? At first glance, this might seem like a straightforward problem, but it is a surprisingly complex one and has become more so in the last 25 years. This is because almost all archaeological information is now created in digital form and includes an ever expanding range of data generation and recording techniques; databases, text documents, images, videos, sound recording, aerial photographs, satellite images, laser scanning, digital mapping, bathymetry, three-dimensional models, etc. It is often very surprising to people, even experienced archaeologists, to discover that with all this new technology, and sometimes because of it, the data created is really very fragile and requires a significant amount of management to keep safe. It really is counter-intuitive that digital data requires more direct intervention and management to keep safe and usable over time than a traditional paper record. Having said that, we have all experienced hardware formats going out of date and becoming virtually extinct. Remember floppy discs, the minidisk and the audio cassette? We have also had experience of corrupt files, software changing such that old documents or data files are no longer read by the latest software, or software and file formats being restricted to particular devices; just think of the iPhone.

These are trivial examples, but they represent exactly the kinds of issues that face all kinds of digital data. Very often, without intervention, important data, created say just ten years ago, is simply not accessible or usable. Already we at the ADS know of significant amounts of important and irreplaceable archaeological data that has simply vanished due to corruption, hardware, software or metadata issues (i.e. we have the data, but we do not know what it relates to or how to make sense of it!). Obviously there are approaches to managing this data to keep it safe over long periods of time; the most frequently cited is the Open Archival Information System (OAIS, `http://tinyurl.com/7vmwu7d`). This comprises hundreds of pages of guidance and good practice, but is only a reference model. Actual digital preservation practice based on OAIS can be enormously complex and expensive – remember in archaeology it has to deal with literally hundreds of file types, from hundreds of types of devices, software and practice. It is simply not possible for an individual or even an individual project to manage the whole process of digital preservation without assistance. This of course is where the ADS come in. The ADS is a digital archive with two main objectives: firstly to provide a safe place for those interested in keeping the results of their archaeological work available to others in the long term; and secondly, but just as importantly, to explore new ways of making all these exciting results available, findable and usable to anyone and everyone over the internet.

Sharing of information and particularly engagement with broader non-academic audiences, community groups and the public is increasingly becoming a stated aim of archaeological projects. There are lots of reasons for this, some based on a genuine belief in the importance of engagement and a desire to democratise the discipline. Others are a reaction to the demands of the REF and the expectation that academic research can demonstrate 'impact' outside the confines of academia. Irrespective of the motivation, having your research outputs openly available online, via a digital archive is a demonstrably efficient route to both reaching wider audiences and creating opportunities to enhance the impact of your research. It is very easy for a single project web site to get lost in the billions pages that comprise the web. It can be a time

consuming and difficult task to keep it on line and keep its contents usable. This is a further advantage of depositing material with a digital archive such as the ADS, once it is deposited its management, exposure and usability all become our responsibility. The issue of exposure is an important one, because isolated datasets represent a problem both for the creator and for potential users. Breaking down barriers between isolated archaeological datasets remains a constant challenge for the ADS, and great efforts have been made by us via a number of major European research and infrastructure projects to find ways of meaningfully aggregating datasets, drawing them together in a way that makes it easier for the researcher to find what they are looking for and have confidence in the results of their queries and searches. The most obvious benefit to archaeology has been that digital sources, such as national and local monument inventories and digital repositories containing significant resources need no longer be hermetic entities relating to specific geographical, period or subject boundaries, but can make their content available as part of a larger aggregated dataset. Thus researchers and cultural heritage managers are enabled to retrieve data on their area of interest from a single point of access without being constrained by arbitrary, technical, political or regional boundaries.

Figure 1 – Who Remembers Floppy Discs? (Image Copyright – Archaeology Data Service)

What is the result of all this work at the ADS? For the data creator, peace of mind and the benefits of having their data publicly exposed and integrated in to broad data discovery and research infrastructure. For the user, the ADS offers access to over 1.1 million archaeological records and around 450 individual archives, some of these, such as the Channel Tunnel Rail Link archive containing data from hundreds of individual excavations and interventions. One of these

archives, the Grey Literature Library now holds nearly 17,000 geo-referenced reports on archaeological events covering every region, time period and type of archaeology. Only a few years ago none of this unpublished grey literature was easily available to researchers and as such a vast potential research resource lay untapped. The ADS's work within this area has been recognised recently with a short listing for the BAA Innovation award, we will hear the results of this in July. All of the ADS records and holdings are carefully managed, categorised and made discoverable via sophisticated search mechanisms such as ArchSearch. This tool utilises both the latest online mapping technology and an advanced faceted classification browsing system. Large portions of this data are also shared with other portals, such as Europeana (`http://www.europeana.eu/`), to ensure maximum exposure to the resources we hold. We know this exposure pays off as we keep careful track of our user base and usage statistics. We have hundreds of thousands of users per year from all over the world, downloading around a million pages and more than ten gigabytes of data every month. Given that latest figures put the size of the academic, student and commercial archaeological sector in the UK at around six thousand people, this level of usage represents a major success. Keep your eyes peeled also for a new JISC funded report due in 2013 on the economic value of the ADS to the whole archaeology sector (`http://tinyurl.com/8y8txma`).

We have looked above at the basic responsibility that all archaeologists have to ensure their important work is archived. We have looked at the difficulties this entails, from the fragility of the data to the complexity of the technical solutions in digital preservation. Finally we covered the advantages in terms of academic profile, long term community and public outreach, impact and engagement as well as the significance of the archive for theorising both the profession and the past. I hope you will agree that not only is the ADS an important organisation in that it addresses these issues on behalf of the profession in the UK, but that it is great to have it in York where its work is important for the reputation of the whole department.

So, finally returning to the question posed in the first paragraph, what is the purpose of archaeology? Whatever answer you might settle on, whatever answer satisfies you intellectually, it is likely that sharing your passion for the subject with others is a part of that answer. Think of digital archiving and the ADS as a service there to help you share your passion, your data, your thoughts and your ideas, not just for today, but for a long time to come.

If you would like to keep up to date on the latest from the ADS you can follow @ADS_Update and also our ADS Facebook page, or keep an eye on our blog: `http://tinyurl.com/848tw6n`