Excavation and Monument Data SIG

Working document

V2, July 2014, prepared by AIAC, with input by SIG members

Introduction

When the general public thinks of 'archaeology' the first image that springs to mind is that of excavation (preferably on a sunlit field), while the second is a major monument like Pompeii or the Acropolis. The success of a European infrastructure will thus almost inevitably be judged on how well it answers the twin questions of access to information about both of these categories, as well as the related category of field-survey data. All three of these are complex, with multiple data sets potentially attached to each single individual. For excavation these would be stratigraphy, description, images, finds, GIS data, geophysical data, environmental data , radiocarbon data; for monuments physical description, location data, images, access data, conservation data; for field surveys descriptions, images and finds data for individual sites, and GIS data, by period, material and site type for the survey as a whole.

Access to these data vary, although the respondents to the ARIADNE Stakeholder survey agree that they are generally abysmal. In the case of excavation data, a report on the site in .doc or .pdf format may be available, but the finds data is unlikely to be found in a database format, and is even less likely to be online. The three categories will now be

discussed separately, introducing their problems and the current state of play in various European countries.

Excavations

Discovery: finding sites

At the start of any search for the purposes of research the primary query would be what excavations have taken place that answer my search criteria? (Roman villa, Bronze-Age sanctuary, medieval cemetery, sites in Hertfordshire). In some European countries – Britain, Bulgaria, Romania, Italy, and Holland through websites such as Oasis (ADS, UK), DANS EASY (NL) and Fasti Online (AIAC) – this information is at least partially available for sites excavated after c. 2000. Oasis (http://oasis.ac.uk) provides an online index to a large amount of reports (in pdf/a format) and other material (images, spreadsheets, etc.) of archaeological fieldwork in England, Scotland and Wales. The digital archiving system DANS EASY (https://easy.dans.knaw.nl) includes the e-depot for Dutch archaeology (EDNA), which provides comprehensive coverage of surveys and excavations in Holland. An initial search reveals a summary record: further data is available to registered users, generally in pdf format. In January 2014, 21,500 data objects were archived: 18,500 archaeological reports and 3,000 datasets which consist of photos, GIS, data-tables, drawings etc. In Sweden, SND is a digital repository and works in a similar way as ADS and DANS. Data deposited at SND is made available online either directly or by ordering data (depends on the data). The data is presented in a catalogue with the metadata extracted from the data files or produced by SND who manages and documents the data to ensure long-term preservation, availability and reuse. The documentation is compiled according to the international standard DDI (Data Documentation Initiative, xml-based and exportable). SND also provides Persistent Identifiers (DOI) to each dataset. Fasti Online only furnishes a short summary of each campaign on an excavation in one of the member countries (principally Italy, Bulgaria, Romania, Macedonia, and Albania). There are c. 3150 excavations registered to date, with around 15,000 records. However, the metadata available on the Fasti site means that the director or institution responsible can be contacted for further information. All these databases are accessible through search terms, keywords and a mapping interface, which is bound to be helpful to planning officials as well as those interested in archaeology. A much newer site is that of INRAP, also with a map interface. It now contains records of several hundred rescue excavations carried out by INRAP in France (http://www.inrap.fr/archeologie-preventive/Sites-archeologiques). These are searchable by a by type and period, though not by keyword. The information available beyond a brief summary includes articles, videos and in some cases recordings: finds data and detailed reports are not included although here again could presumably be made available on application to the institution. This site is reserved exclusively for those rescue sites excavated by INRAP, and complements the site of Archéologie de la France Infos (http://www.revues-gallia.cnrs.fr/spip.php?rubrique19), run by the journal *Gallia*, which records various excavations carried out by universities and local associations, and is searchable, though has no map interface. In Greece, the British School in Athens and the École Française d'Athènes collaborate on the production of Archaeology in Greece Online (http://www.chronique.efa.gr/index.php/). Resembling Fasti in its summary reports

(although these are very short indeed), the site so far contains a few dozen sites, searchable in various ways, including a map interface

Outside of these projects there are a few very intensive projects which produce exhaustive coverage of all the excavations in their respective cities. Foremost is the Museum of London's LAARC archive (http://archive.museumoflondon.org.uk/laarc/catalogue/) which provides an online resource for the study of all archaeological research on in the city, and may be the world's largest archaeological archive. However Rome's SITAr and Pisa's MAPPA also aim to provide comprehensive records of all excavations which have taken place in their respective towns. Both projects are linked to ARIADNE as associated partners, but they are not yet available to the public. For Vienna a government site maps all monuments and archaeological sites in the city (http://www.wien.gv.at/archaeologie/), with date of excavation, site codes and bibliography. Apart from these (and other projects which we have undoubtedly missed) a member of the public looking for excavation reports from a given country will have a difficult job knowing where to start.

Some resource discovery services specializes in specific types of sites. There are for example open access online services on archaeological evidence of human remains. The Mapping Death database (http://www.mappingdeathdb.ie/login) is an open access site that provides information on graves in Ireland 1st – 8th AD: location, descriptions, bibliography, scientific dating. Data can be queried via a map interface and text, allowing to search for burials according to periods, burial types etc. or also specific features such as trauma, decapitation.

The French project EMA - L'Enfant et la mort dans l'Antiquité (Death and the Child in Antiquity) http://www.mae.u-paris10.fr/ema/ published an open access online database for children's graves from the ancient Greek and Roman world (beginning of the 1st Millennium BC to the end of Antiquity). The database holds records of 4.000 child graves and information on topography, location of the grave in funerary or other areas, markers, graves, containers related to each individual and goods. More than 90 % of the documentation is available for consultation, after the consent of each of the concerned contributors. The languages of both database and website are French and English.

Long-term storage and documentation of excavation data

Archiving and storage of excavation documentation is addressed by only two of the partners, **ADS** and **DANS**. This is in part because it represents a far more serious undertaking, including questions of sustainability, access, copyright and data interoperability. The solution adopted by ADS requires the excavator to provide metadata on the components of the submission, which tend to include full reports of most aspects of the excavation. The finds reports are not always in a database format, making their use and linking much more problematic. DANS EASY adopts a similar policy, giving persistent modifiers to the documentation received. These are the only national archives of excavation data currently available: other excavation data is stored in the institutions that carried out the research, dispersed through university departments and very seldom available online.

Two American projects, the Digital Archaeological Record (http://core.tdar.org) and Open Context (http://opencontext.org), guarantee long-term, sustainable storage of project data in return for a fee. Data may be presented as CSV files with metadata, which helps retrieval, as well as storage of images and project reports. tDAR is not easy to browse, and

lacks map interfaces except at the level of the individual site. It is clear that one would have to know that a site of interest existed in order to access any data related to it, although at the level of an individual site the data may be very complete. Open Context contains large amounts of data from a range of sites, some of them in Europe but most in the Middle East. It is fairly easy to search the site and to retrieve data from it.

A minority of sites serve their own excavation data (stratigraphy, finds catalogues and so on) on dedicated websites, usually housed on university servers. Examples are the Via Gabina villas project (http://viagabina.rice.edu) and the excavations at Cosa between 1990 and 1997 (http://www.press.umich.edu/special/cosa/). The Via Gabina site is essentially the whole publication, including the catalogues in database format, while the latter complements a published volume, providing the stratigraphic and anthropological data for the volume, as well as tables of sherds. Both of these sites are in a very simple html format, and present few problems of sustainability, in spite of the general lack of metadata. A more dynamic solution is that of the Prescot street excavations in London (http://www.lparchaeology.com/prescot/) which serves live data via the ARK system. The most elaborate excavation database online remains that of Çatalhöyük, http://www.catalhoyuk.com/database/catal/Search.asp. This sort of solution seems best-applicable to large scale research excavations that are capable of convincing institutional servers to house them. However, universities have become increasingly wary of such projects, because of the apparent risks posed by interactive websites, and issues of long-term sustainability. Further, although data may be available on such sites, it is difficult to link it to other such data, although the individual catalogues can generally be downloaded in CSV or Excel formats.

Although paper documentation of excavations has the disadvantages of inaccessibility in most cases, the vast majority of excavation records that were 'born digital' are at risk. The rapid changes in technology inherent in digital archiving it can be difficult for museums and universities to curate, and the link between excavation and object can thus easily be lost. Complex relational databases built for excavations a decade ago on Microsoft Access can now no longer be read without a computer still running Access 6: those on more obscure proprietorial databases are entirely inaccessible. Thus some form of sustainable deposit is a major necessity.

Monuments

Sites and Monuments Records

National sites and monuments records are generally extensively catalogued by governmental institutions (for the UK the ordinance survey and English Heritage at the national level, and then at the county level). In Sweden, for instance, an online database "Fornsök" (http://www.fmis.raa.se, in Swedish only) created by the Swedish National Heritage provides basic identification of sites, often with geo-references, sometimes with fuller descriptions and links to images, management information and other resources including GIS. Such resources are generally designed to support the management of archaeological sites, with different levels of protection (from recognition on world heritage registers, through national listing, local listing to simple identification). They provide a primary source for desk-top research and other forms of archaeological investigation carried out in advance of development. They also provide a resource for example for identifying

areas of archaeological potential, for landscape archaeology and in understanding the context and settings of archaeological sites.

Many of these databases are available online (an example is the Archaeological Survey of Ireland (http://www.archaeology.ie/ArchaeologicalSurveyofIreland/). A more ambitious project is the CARARE project, coordinated by **MDR**, which focused on aggregating archaeological and architectural sites and monuments for the Europeana initiative. The remit of the CARARE project concentrated on the infrastructure needed to support the aggregation of inventories (amongst other resources) created in different countries according to slightly different standards and in many languages. http://www.carare.eu. The project produced a prototype map tool (http://carare.eculturelab.eu/Carare50m/Map.html) which provides the locations of sites ranging from major monuments to excavations and artefacts on a GIS map base. The site integrates several monument inventories and other resources (including data from **DAI**'s ARACHNE database amongst others) and the information is served it on a single map where it can be browsed using keyword searches or by using route-planning tools.

A remarkable experiment in the combination of various forms of archaeological data on a single internet site is the Getty-financed MEGA Jordan (http://megajordan.org/) which provides GIS documentation of sites ranging from major monuments to excavations to field-survey sites. More of a sites and monuments record than an excavation database, it is still interesting in its integration of this information and the fact that it serves it on a single map: we will return to this resource at the end.

Survey Archaeology

A major change in the archaeological focus in the last quarter of the last century was the growth of archaeological field survey, studying large tracts of land through intensive field-walking, combined with aerial photography and occasional geophysics. These projects, rather than providing in-depth information about a single site, created a more-or-less precise view of a landscape, with the degree of detail depending on the intensity of the survey. As time has passed many of these projects have been published on paper, some as volumes, others a journal articles. In the case of final publication there are also catalogues of materials. A favorite thesis in Italian universities, in Italy alone they number in the hundreds. Many projects were later retrofitted to GIS format: a case in point is the decades-long Tiber Valley Survey of the British School in Rome (an associate of ARIADNE). However, the prospects for on-line availability of these projects, where the aggregate would be substantially more than the individual parts, are still inexistent. Indeed, the Tiber Valley Survey plans long-term sustainability through deposit with **ADS**, with the consequent loss of the GIS component.

In Bulgaria surveys since the 1990's are recorded on an online information system "Archaeological map of Bulgaria" (AIS AKB). All archaeologists are obliged to fill in information about newly registered sites and monuments as a result of their annual fieldwork. However, access to the database is restricted and protected by a regulation of the Ministry of Culture who is its actual owner. **NIAM-BAS** is responsible for its protection and maintenance. Despite its name, it is not GIS-based system, although the intention is to create a mapping interface.

A very few surveys are fully available online: an example is the Jerba Project, held on a university server (http://www.sas.upenn.edu/jerba/index). The map can be searched by site type and date,[1] and various catalogues, such as architectural fragments, are available. A discovery tool for identifying 317 survey projects is represented by Mediterranean Archaeology GIS (MAGIS: http://cgma.depauw.edu/MAGIS/). It is more or less limited to Anglo Saxon projects, never having reached the national universities and foreign schools that are the generators of many of these projects, and, unfunded, has not been updated for a few years. Survey archaeology is thus the least available on line of all major forms of archaeological research. A preliminary investigation is underway with the British and Dutch Schools at Rome to study the feasibility of a common interface for field survey research in Italy.

Needs and Challenges

Discovery

While the scope of ARIADNE's remit concentrates on the creation of an infrastructure that will make cross-platform searching possible, in the hierarchy of needs the aspect of discovery remains a major challenge. Digital storage of excavation records is certainly vital, but below that there are, of course, thousands of excavations whose records were born on paper and have never been digitized. These, however, are generally accessible in some form (excepting those which are thrown out with the death of the excavator). It is the existence of the excavation that needs to be recorded in some format easily available online, complete with metadata that will give some idea where and by whom the archives are housed. The very simplicity of the Fasti and INRAP records makes them quick and easy to fill in, though the legal obligation that requires that excavations be reported in Italy and Romania makes their records far more comprehensive. In the same way discovery of survey projects would be a major step forward towards their exploitation for research.

.

Sustainable storage and metadata

There is no doubt that the amount of work that may be involved in the deposit of digital excavation records represents a major obstacle, perhaps even more important than the cost of the deposit, although that may not be insignificant, and is rarely considered in excavation budgets. Without the mapping of the data no cross platform searching is possible, but getting excavators to budget this into their time and resources may represent a considerable challenge. The question of long-term sustainability is also a major challenge: even such rock-solid institutions as ADS are occasionally threated with de-funding. As one of the stakeholders commented, maintenance is not politically popular, nor is funding easily available for long-term projects (witness the collapse of MAGIS). The price of availability is constant, and reliable funding.

[1] The map is currently malfunctioning.

How can ARIADNE help?

The common denominator of excavations, monuments and survey projects is the existence of point data that can be mapped and viewed online. However, national monument inventories are mapped using national mapping systems, integration on a European or world map requires conversion of the national coordinates to an international spatial reference system such as WGS84. Providing tools, which support and enable this conversion, will enable the development of a GIS-based European portal to sites, excavations, monuments and buildings.

A second issue is the creation of international standards for the documentation of excavations and monuments so as to render it transparent and comparable. Free access to tools, particularly for data mapping, to make it easy to comply with these standards will be important, as will be offering the means and guidance to archaeologists to deposit their digital records. The sustainability of digital datasets must also be high on the agenda.

Once these problems are resolved, however, it would be possible to create a complex map with a GIS of sites that could be enhanced by differentiation of the points (survey sites, excavation sites, listed monuments) and layers that could be turned on and off. Such a map would make discovery of broad categories of sites in a given landscape far easier, while allowing the user to understand at a glance what sort of further information might be available: MEGA Jordan is here the model for the geographic display of disparate types of site. Data points would then create links to the site from which they were drawn, using the Pelagios/Pleiades systems which already ties them together, and thus further research could be carried out. Once such a structure is in existence individual institutions could contribute datasets (this would be particularly true of field survey data) which could then be served on institutional computers. The result would be a multi-sourced cloud of data, from which a variety of institutional datasets could be discovered. The institutions would then become stakeholders in the care and maintenance of the collective site.

Of course, such a project has its utopian aspects, and there are parts of it that might best be built from the ground up. On a larger scale, the creation of such a resource would be a preliminary result that would allow both professionals and the general public to gain some idea of what is out there. Rather than a traditional portal, which always implies some kind of list of sites whose purpose is more or less opaque, a GIS-based European portal to sites, excavations and monuments would provide on an international scale the sort of resource that ADS and DANS offer at a national one.

The archaeological community at large is often unaware of existing online resources. A major need would be to provide an up to date collection of online resources on most archaeological topics. For example, all online resources related to burial archaeology. A first step for a research infrastructure for archaeologists will be to bring together t a list of links, even those that are not ‘integrated’. The creation of such a service could be ongoing, structured so that new resources could be added by those providing them. Thus ARIADNE would serve to link even those projects that, without metadata, could not be entered into the new infrastructure