

The Transatlantic Archaeology Gateway: Bridging the Digital Ocean

Stuart Jeffrey, Lei Xia, Julian Richards, Jon Bateman
University of York, UK. stuart.jeffrey@york.ac.uk

Keith Kintigh, Francis Pierce-McManamon, Adam Brin
Arizona State University, USA. kintigh@asu.edu

Abstract:

The Archaeology Data Service (ADS) founded in 1996, is a national repository for digital data from the UK historic environment sector. In the USA Digital Antiquity and its digital repository, the Digital Archaeological Record (tDAR), was established at Arizona State University in 2008. Both organisations are acutely aware of the problem of 'siloes' data, i.e. data only discoverable and/or accessible from a single location, and actively research how to make their holdings widely available. Funded by Joint Information Systems Committee (UK) and the National Endowment for the Humanities (USA), the Transatlantic Archaeology Gateway (TAG) has developed services and a search interface, for cross-searching between ADS and tDAR at archive collection level. This project is the first example of this kind of technical cooperation between digital archaeological archives based in Europe and the USA and has significant lessons for the development of cross-searching and data sharing between such geographically remote archives.

Key Words: *Digital Archiving, Service Oriented Architecture, Cross-Searching, ADS, tDAR*

Introduction

The use of ICT for teaching and research has become well established in archaeological practice on both sides of the Atlantic and discipline of archaeology has often lead the way in adopting these technologies when compared to other humanities. This is not least because modern archaeological practice results in the routine creation of large volumes of primary digital data. This data is generated not only by the use of sophisticated recording techniques such as geophysical and bathymetric survey, laser scanning and photogrammetry, but also by the extensive use of databases and rich media as standard excavation recording methodologies. This results in complex datasets in a variety of formats, including structured and unstructured text, spreadsheets

and databases, still and moving images, CAD, GIS, landscape and object-scale 3D scans, and virtual reality models. Much of this primary data is born digital as standard and as the only record of unrepeatable fieldwork it is essential that these data are preserved, for re-use and re-interpretation.

In the UK the Archaeology Data Service (ADS 2011) has grown to become the leading national repository for digital data from the UK historic environment sector, cross-cutting the academic and public and private sectors (Richards 2008). In the USA, more recent developments in creating a national archival infrastructure saw, in December 2008, the Digital Antiquity initiative and its digital repository, the Digital Archaeological record (McManamon & Kintigh 2010; tDAR 2011) established at Arizona

State University. With generous funding from the Andrew W Mellon Foundation, Digital Antiquity's scope also covers the whole discipline and includes archaeology from the private, governmental, and academic sectors. Both the ADS and Digital Antiquity have a clear understanding that their archival role is underpinned by a need to provide as access to their data set holdings to as broad and audience as possible. Integral to that role is the belief that infrastructures entirely locked within institutional or national boundaries are detrimental to wider research processes in archaeology. These boundaries often mean that researchers cannot easily discover, let alone access, the totality of datasets relevant to their work without the potentially onerous process of independently discovering every possible repository and engaging is an exhaustive search process within each. In an ideal world each repository of archaeological datasets could serve up details of their holdings via a shared gateway and a scholar would be able to use this point of access to discover, cross-search and subsequently access all relevant datasets. Archaeological research is, by definition, international. For most of the human past modern political boundaries were irrelevant. The 'big questions', including hominid evolution, the development of agriculture and sedentism, the growth of complex societies and urbanism, human impact on the environment and so on, all transcend the modern political map. In order to pursue an understanding of major changes and to investigate whether these follow independent evolutionary trajectories or stem from diffusion or migration of people and ideas, it is necessary to have access to data sets which cross countries and continents.

This paper describes some of the key results of a major collaboration between the ADS and tDAR called the Transatlantic Archaeology Gateway (TAG 2011) funded by the JISC/NEH Transatlantic Digitisation Collaboration Grants. The TAG project builds upon NSF cyberinfrastructure initiatives in the USA and

the eScience programme in the UK to provide a shared approach to resource discovery and data mining. Running from October 2009 up to April 2011, the primary aim of the TAG project has been to develop tools for transatlantic cross-searching and semantic interoperability between ADS and tDAR. TAG has developed interoperability between the USA and UK at two levels. The first stage has been to create an infrastructure to enable basic cross-search of Dublin Core compatible metadata records for digital resources covering the archaeology of the USA and UK. This has built on earlier work on the EU-funded ARENA and ARENA2 projects which demonstrated such an approach was achievable within Europe when dealing with monument inventory data. Nonetheless, mapping European to North American metadata schemes offered some real challenges, particularly with regard to periodization and subject type. The second stage of TAG was an attempt to develop a much deeper and richer level of cross-searching for a specific subset of archived data: faunal data from North America and Europe. This sub-discipline was chosen as there is a relatively high level of agreement over basic classifications. The provision of such deep data mining would be ground-breaking and the TAG project's attempts to achieve this provide useful insights into the challenges faced when trying to draw together existing and heterogeneous datasets even when their topics and content are ostensibly so similar.

Background and Architecture

The TAG project antecedents were in two European Union funded projects that tackled a similar problem, aggregating monument inventories from a number of countries for cross-searching. These projects, The Archaeological Records of Europe Networked Access (ARENA) 1 and 2 explored some of the technical approaches likely to bear fruit for a project extending searches across the Atlantic and across archival datasets.

The archaeological records of Europe Networked Access (ARENA) projects 1 and 2

The original ARENA was a Culture 2000 project completed in November 2004. The ARENA web-based portal was completed in 2004 and launched at the European Association of Archaeologists conference at Lyon. The ARENA search portal, based on Z39.50 and OAI harvesting, facilitated searches for archaeological sites and monuments inventories from six European countries. This project is extensively documented in publications by Richards, Kenny, Kilbride (2005) and Waller (2005) and there is a special edition of the journal *Internet Archaeology* (Issue 18) almost entirely dedicated to the project and its implications for academic archaeology in Europe (*Internet Archaeology* 2005).

ARENA2, funded as part of the Preparing DARIAH project, changed the technical approach of ARENA to one based on a Service-Oriented Architecture (SOA) over selected partner data centres to demonstrate the viability of using a SOA approach rather than the original Z39.50 and OAI metadata harvesting approach. A SOA approach differs from metadata harvesting approaches in that it allows direct, live access to remote databases making the most current data available to be queried and minimising harvesting and data management on the part of the aggregator:

*“The OASIS SOA Reference Model group defines **Service Oriented Architecture** as a paradigm for organizing and utilizing distributed capabilities that may be under the control of different ownership domains. It provides a uniform means to offer, discover, interact with and use capabilities to produce desired effects consistent with measurable preconditions and expectations.”* <http://www.oasis-open.org/>

The objective of the ARENA2 demonstrator was to migrate ARENA into a sustainable environment by adding service logic and

exposing its resources as autonomous services in a Service Oriented Architecture. The basic ARENA2 architecture followed the ‘Publish-Find-Bind’ approach. Services publish themselves to a registry as being in accordance with a web service specification. These services are then found and bound to by a client. In this case the specification is the ARENA Gateway Service Specification, the client is the ARENA2 portal, the services are either compliant monument inventory services or ‘wrapped’ services based on legacy protocols such as Z39.50 or OAI PMH and the registry is an instance of a Universal Description Discovery and Integration registry, the ARENA UDDI registry.

The ADS’s experience with the ARENA and ARENA2 projects pointed to the likelihood that a similar general approach to ARENA2 would be the most fruitful for TAG. In essence, this is the creation of a common service specification, registering of services in a registry and the construction of a portal that guides the user in the construction of queries and then sends these to each service target (repository) before aggregating the results into a single form for the user to browse.

The TAG architecture

Although the question of the general approach to the TAG architecture was fairly easy to resolve, the construction of the proposed system fell into a number of distinct phases. A TAG service specification had to be created and agreed between the UK and US partners, and a Universal Description, Discovery, and Integration (UDDI) registry had to be implemented to allow the services to be registered. Each partner was then required to ensure that they were able to serve their archive metadata as a service adhering to the service specification and allow querying. Once these services were in place and registered it was then necessary to create a service cross-search

and result aggregation interface. Once this was in place, allowing access to the archive services of the ADS and tDAR, the original aim was to repeat each step process in order to allow deep cross-searching (i.e. record level) of the faunal remains datasets held in each archive. The sections below describe each stage in the process in more detail:

Agree and extended/amended service specification

Integral to creating a SOA implementation of the TAG service is the specification of the specific modes by which the services should communicate. Written in Web Service Description Language (WSDL) the TAG specification details how requests to the target services will be structured and how the results are expected to be delivered from the target service (repository) back to the result aggregator. Both requests and results are sent via the hypertext transfer protocol (http) using XML. However, it is necessary for both the portal creating the query and the target service to know the structure of this XML. Drawing on the experience gained during the construction of the ADS's faceted classification browsing tool (the JISC funded Archaeotools project) (Jeffrey et al. 2009) consensus between the UK and US partners emerged around the continuation of the 'Where, What, When' query structure adopted in the two ARENA projects and the ADS ArchSearch browser. The most significant reason for this is that, as with monument inventories, virtually all archives have metadata at collections level that covers these elements in a single record. Importantly, these are the most frequently used search criteria for archaeologists academic or otherwise (Kenny et al. 2005).

The TAG Service Specification itself consists of a WSDL document. This is an XML document describing a Web Service - the location of the service and the operations (or methods) the service exposes. The TAG Service Schema is

referenced by the TAG Service WSDL document, defining the specific methods TAG Service can implement: The UDDI enabled TAG Registry further requires a document representing the service data structure, known as a tModel. The tModel is an abstraction for a technical specification of a service type, organizing the service type's information and making it accessible in the registry database. The TAG Service tModel describes a Gateway Service by including a pointer to the Gateway Service WSDL document. In this way the TAG Registry can process a client query to return a list of available services complete with the necessary information to enable a client application to bind to.

The Where, What and When facets

As mentioned above a 'What, Where, When' structure was settled on based on the experience of the ADS and tDAR in designing search mechanisms for a diverse user base that balances simplicity with powerful and meaningful search functionality. Having decided on these key facets further discussion was required between the partners to detail what, in practice, these facets would relate to and which controlled word lists, thesauri or data structures would be used to represent each facet in the interface and in the subsequent query.

In order to create the service specification for each of the 'Where, What, When' elements a universally agreed and available controlled list had to be selected to which each of the data sources in the UK and the US could be meaningfully mapped.

WHERE – the search interface would require a geospatial selection method that did not rely on local coordinate systems in either the UK or in the US. This means that to either plot or to select via map archives from either, or both, countries their spatial component would have to be represented in a universal coordinate

system. The Latitude and Longitude WGS84 coordinate system was selected as the location coordinate system for exactly this reason. This required data held in the ADS using local coordinate systems to (e.g. OSGB 36) to be mapped to WGS84.

WHAT – This is the facet indicating what type of archaeological site or sites the archive is most concerned with. Of course the ‘What’ of an archive can also relate to what types of information the archive actually contains, e.g. a finds database, GIS files and so on. For initial resource discovery though the first type of ‘What’ is considered most pertinent. The best developed thesaurus for monuments in the UK (and arguably Europe) is the English Heritage Thesaurus of Monument Types (see English Heritage’s MIDAS standard, English Heritage 2011). This a poly-hierarchical thesaurus developed over a long period of time and with entries covering all, or nearly all, site and monument types occurring in England. Versions of the TMT also exist in extended form to cover additional sites and monument types in Scotland and Wales. However, in a US context a number site and/or monument types are not likely to be present in the thesauri, or even where terms are shared, the actual local usage in the UK and the US might not actually be identical. A process of negotiation was engaged in where by each level of the thesauri was looked at until the solution of mapping of archive records to the 16 ‘top-level’ terms was arrived at. These terms are not in fact site or monument types at all, but they are functional groupings into which sites and monuments can be classified. It was agreed that all sites and monuments on both sides of the Atlantic could be classified into one or more of these classes.

Relevant TMT top level terms:

- “AGRICULTURE_AND_SUBSISTENCE”
- “CIVIL”
- “COMMEMORATIVE”

- “COMMERCIAL”
- “COMMUNICATIONS”
- “DEFENCE”
- “DOMESTIC”
- “EDUCATION”
- “GARDENS_PARKS_AND_URBAN_SPACES”
- “HEALTH_AND_WELFARE”
- “INDUSTRIAL”
- “MARITIME”
- “RECREATIONAL”
- “RELIGIOUS_RITUAL_AND_FUNERARY”
- “TRANSPORT”
- “WATER_SUPPLY_AND_DRAINAGE”

WHEN – This facet relates to what archaeological time periods the archive relates to. As with the ‘What’ schema alternative information could be covered by the ‘When’ facet, such as when the archaeological recording event was carried out or when the site was first recorded. In the UK the Forum for Information Standards in Heritage (FISH) maintains the Manual and Data Standard for Monument Inventories (MIDAS) which has at its heart a number of controlled vocabularies. For ARENA and ARENA 2 this list proved adequate for mapping most European descriptive periods, although cultural period values, such as ‘ROMAN’, vary enormously from region to region in the absolute date ranges that they represents and does not exist at all in some regions, a differential mapping to cover this variation was established in the ARENA programme. For the TAG project it was clear from the outset that the descriptive values used by MIDAS and those in common usage in the US would not map meaningfully.

The MIDAS period description list, relevant terms:

- “PALAEO LITHIC”
- “MESOLITHIC”
- “NEOLITHIC”
- “BRONZE_AGE”
- “IRON_AGE”
- “ROMAN”
- “EARLY_MEDIEVAL”
- “MEDIEVAL”
- “POST_MEDIEVAL”
- “MODERN”

Examples of US descriptive periods.

- “EARLY PEOPLING”
- “PALEO-INDIAN”
- “ARCHAIC”
- “VARIOUS REGIONAL TRADITIONS” i.e. cultural terms would normally be used here.
- “CONTACT”
- “HISTORIC”

Fortunately normal practice in the US is to specify absolute dates (i.e. numeric dates BC/AD, CE or BP) for a site and this data becomes metadata for archives deposited in tDAR. In the UK period description terms are the norm, however these can generally be mapped to absolute date ranges with some confidence. Although, as already pointed out, the absolute dates relating to a such cultural terms are necessarily contentious and at the very least periodical re-mapping would be required to make a search on an absolute date meaningful when the underlying data was in fact a cultural term. The ADS enhanced its own collection level metadata with the addition of absolute date representations of its period terms. This now meant that for the ‘When’ facet absolute dates could be used to cross-search meaningful between tDAR archives and those at the ADS.

Creating UK/US resource inventory services according to service specification

For both partners the specification of the TAG service structure and the ways in which it would generate queries against the three facets detailed above triggered the need for three things; first the construction of the ADS and the tDAR target service and their registration; second, a way of managing the service logic by converting the TAG XML queries to queries of the underlying database system (in e.g. SQL) and then converting the results to XML in a form expected by the portal. Finally each partner also had to engage in a metadata analysis exercise to ensure that the ‘What, Where, When’ data they held for each archive was indeed in a form that followed the TMT top level terms, absolute dates and WGS84 coordinates. For the ADS this meant conversion of place name data or OSGB36 coordinates to WGS84 and the mapping of all the period descriptors to absolute data ranges. For tDAR the challenge was to make sure that each archive had metadata adhering to the top level TMT functional groupings.

Creating and Running a Query

With the existence of the TAG service specification and the construction of the archive discovery services themselves, one at the ADS and one at tDAR, the next stage was the creation of a portal via which a user can construct a query that is subsequently sent to each registered archive discovery service. Following the ‘Where, What, When’ model a portal website was constructed allowing each of these facets to be specified via user-friendly mechanisms. The design approach for the portal was to make it transparent to the user that he was ‘constructing a query’ using a set of tools and that that query, once complete, was sent off to the target services. To this end the search page is split with the top of the page clearly marked ‘query’ and the lower section of the page marked as ‘query builder’. It is in this lower section that the visual tools to guide the



Figure 1. The TAG Portal Interface (TAG Search Interface 2011).

user in selecting a 'What, Where and When' are located. In addition to these facets, and because it is so key to almost all modern search paradigms, a keyword search box is included. The keyword is specified as a query variable in the TAG service specification and each service target decides which are the most appropriate metadata field(s) in their repository to index and search across using this search term.

The Where query

The obvious solution to selecting a geo-spatial area is to use a map-based interface. As discussed above each archive in the US and the UK repositories now has a location expressed in WGS84. By embedding an open layers (openlayers.org) web mapping window the user is given the ability to navigate the map, pan and zoom and also by toggling a radio button an area of the map can be selected as a bounding box search area. In figure X below the coloured area on the map represents the selected search area and the Latitude and Longitude of the bounding box can be seen in the 'query' section of the search page (for full colour image please see the online version of this paper).

It should also be noted that a lot of archaeological work in the 'Old World', such as Classical archaeology, is carried out by US institutions and therefore most likely to be archived in US repositories. This means that a geo-spatial search specifying a non-US Latitude/Longitude bounding box as a search area is still likely to return results where the archive is held in the US. It also means that the search box must be capable of panning well beyond the UK and the US a proportion of archives held in both the ADS and tDAR repositories actually relate to work carried out elsewhere in Europe and the Middle East.

In figures 1 and 2 the open layers window can be seen using Google Map as base mapping, but any available base mapping layer can be drawn into an open layers window.

The What query

The 'What' facet of the query was probably the easiest element to create a user interface for. The simplest and most direct method was to use radio buttons to allow users to select one or more of the functional group search terms. These can also be removed from the query by means of a delete button (a green and white 'X') next to the selected term in the 'query section of the page'.



Figure 2. The map area selection module (Open Layers).

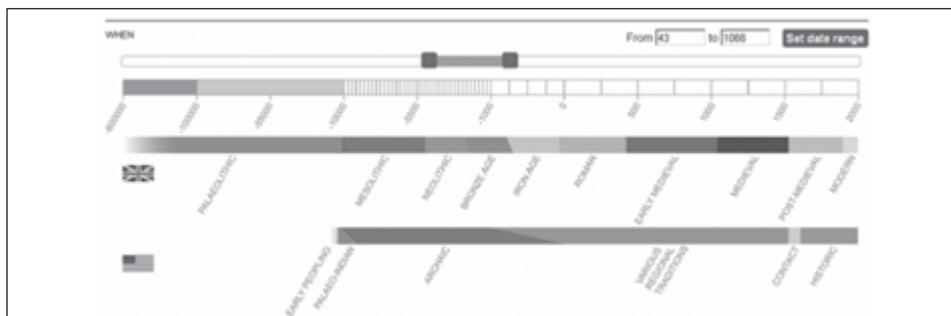


Figure 3. The TAG time slider date range selector and descriptive period guide.

The When query

Unlike both the ARENA projects and the ADS ArchSearch interface which uses a similar approach to the 'What' query selection (i.e. radio buttons), the fact that the underlying data supported searching using absolute dates meant that a more sophisticated visual interface could be employed (Fig. 3).

For this facet the TAG project used a 'timeline' style interface which allows the user to set both the upper and lower bounds of the search in terms of years. The easiest way of doing this is actually a simple pair of 'From – To' data entry boxes. These are actually included in the interface, however, the ability to use the slider on the timeline in association with the descriptive period guide images below them is probably a more intuitive way of setting the desired date range. The time slider employed is in fact carefully scaled so that the time periods with more detailed gradations (say from the Mesolithic/Archaic periods onwards) get a proportionally larger length of the time slider that much longer, but less gradated periods such as the Palaeolithic.

Once the user has defined their areas of interest by selecting a bounding box, a subject or subjects and a time period (perhaps narrowed with a keyword) the 'SEARCH' button is used and the query is sent to both the ADS archive service and the tDAR service. Locally within each repository the XML is then translated into a database and the results in turn converted back to XML and sent to the portal for display. In this portal

each results set is so similarly structured that they share the same XSLT, transforming them for display in HTML. Only the top five results from each repository are actually displayed (although a numeric indication of the total number of relevant results is shown). The user can page through each result set by selecting the 'See More Results' link which opens a pagable version of the results in a separate browser window. Whilst the results displayed in the portal are themselves metadata records, included in each is a hyperlink. This means that when the user has browsed the returned results they can simply click on the desired record to be taken directly to either tDAR or to the ADS. Here the full repository record and archive material are available for download. In this way a user can, with three or four clicks, search remote digital repositories on either side of the Atlantic simultaneously and find downloadable datasets for immediate use in their research.

Although it is clear that the TAG portal is highly usable and represents an excellent starting point for the development of more complex, deeper and broader searching between cooperating archives in this and other domains, time constraints meant that some usability features which we would have like to see implemented were not completed. This is particularly true of the limitations imposed by having only a single search term allowed for the geo-spatial interface. Whilst straightforward to implement, and excellent for searching within a country it does not make it that straightforward to cross-search the UK and the US together, let alone include other countries. This is because,



Figure 4. The query and results pages showing links to archive collections.

unlike the 'What' and 'When' elements of the query builder only a single instance of the facet can be selected for inclusion in the query. In essence this means that only a single bounding box can be selected at time. A very positive future enhancement of the system would be the ability for the user to select multiple bounding boxes. Similarly it would be helpful for more than one date range to be selected in a single search. The interface design changes required to do this, whilst not trivial, are far from insurmountable but it should be noted that it would also be necessary to change the TAG service specification to allow for multiple search terms and consequently the back-end searches implemented on the target repository systems would have to be updated.

Record Level Cross-Searching

Given the richness and diversity of the datasets held in each TAG repository i.e. the broad range of file formats identified in the introductory section of this paper, it is fair to say that aggregating all the archival material into standard schema to allow cross-searching within each individual collection at record level would be a massively challenging task. However, the TAG project was determined to investigate this possibility for at least one small subset of all the types of data and all the

formats held. Faunal remains datasets are very commonly generated in the post-excavation phase of archaeological projects. These data describe the animal remains recovered during excavation, often in very great detail, covering numbers, species, condition, measurements etc. These data are ideal for recording in databases and have generally been recorded this way by the faunal remains specialists for many years. There are a number of important archaeological questions that could be addressed if these data could be searched at record level across sites and across time periods. Given that there is some level of agreement amongst specialists in the field regarding which data points to collect for faunal remains it was proposed that for the TAG project an agreed generalised search schema could be developed. With such a schema faunal remains databases could be mounted in the UK and US repositories in a way that allowed such cross-searching via an extension of the TAG portal interface.

Creating a Faunal Remains Schema

A workshop was held at the ADS offices at King's Manor, University of York drawing together leading figures in faunal remains analysis in the UK and the US. This workshop's main thrust was an intensive discussion on the topic of what fields (or data points) would comprise an acceptable minimum for faunal experts to search on if they wished to cross-search data within a faunal remains database. Unsurprisingly, there was a lot enthusiasm for the proposed TAG approach amongst the representatives of the community and an in-depth discussion took place on the proposed search schema, especially with regard to the level of detail that would be required to provide meaningful results from the datasets. More surprisingly perhaps the variation in recording techniques between individual specialists proved to be greater than originally thought and the complexity of the search schema therefore grew. This in turn would increase the complexity of the queries required on individual

faunal remains datasets held in their respective repositories. For example, the list of data points desirable to be able to cross search was 35, a high, but manageable number (compare this with the three data points for searching at collections level). However, these required a further 59 qualifiers to cover what was considered the minimum for meaningful cross-searching at record level. For example, when looking for 'gnawed' bones, a simple yes/no/probable would not suffice as many specialists actually record gnaw marks by the likely mark generator, so the qualifier list for 'gnawed' is actually yes: human, large carnivore, small carnivore, rodent, herbivore, omnivore /no/probable.

Ingestion into tDAR

Given the difficulties in creating both a faunal cross-search schema and designing a query construction interface that would be meaningful for users and the time constraints on the TAG project, the project team decided to look for an alternative means of allowing cross-searching at a record level between UK and US faunal remains datasets. The process of ingestion into the US archive, tDAR, actually requires the depositor to make their dataset at a relatively deep level to a common tDAR schema. This means that the depositor has to actively map their data against the tDAR schema with significant potential for archaeozoological data integration (Spielmann and Kintigh 2011). The tDAR schema is rich enough that individual record elements can in fact be cross-searched. The repository itself can accept datasets from anywhere and it was therefore possible to ingest UK based faunal remains datasets into tDAR. This was carried out by the ADS for a number of collections including those held at the ADS (e.g. Cottam, East Yorkshire) and also UK datasets held by the Museum of London (e.g. Spitalfields, London) (MOLAS 2011). The TAG results page was then altered so that where datasets that were suitable for integration were returned by

a query this was made apparent to the user via the results page, this was done simply via an additional link: 'Compare datasets within these results: Click Here'.

Conclusions

The primary aim of the TAG project, to create an architecture that sits atop US and UK based archaeological repositories and allows the cross-searching of their archival holdings, has been achieved. The cross searching process is facilitated by a user friendly interface that guides the user through the process of constructing a 'What, Where, When' query in an intuitive fashion, while still allowing standard 'keyword' searching. The results returned from repositories on either side of the Atlantic are displayed in a browseable and pageable format that makes it simple for the user to drill down into results to reach the online interfaces of the collections of interest, and subsequently to directly download datasets. Substantial progress on TAG's secondary ambition of allowing cross-searching at a record level within a specific type of dataset (faunal remains databases) was also achieved. A valuable consensus in the faunal remains specialist community on both sides of the Atlantic was created and a meaningful and powerful cross-search schema was developed. Whilst this was not directly implemented in the TAG portal, flagging of tDAR ingested faunal datasets that would support cross searching was implemented and users are clearly directed to these from the TAG portal results page. The outcome of the TAG project represents not just a technical accomplishment and a useful resource for archaeological researchers, it also represents a highly successful collaboration between organisations committed to the long term preservation and access of digital archaeological research outputs. There is a fervent hope on both sides of the Atlantic that this level of collaboration not only continues between the ADS and Digital Antiquity, but that it acts as a model for ever growing levels of cooperation amongst these and other

international organisations. In this way artificial barriers to research represented by modern geo-political boundaries can be mitigated against if not entirely broken down.

Bibliography

ADS 2011. "Archaeology Data Service." Accessed June 30. <http://archaeologydataservice.ac.uk>

English Heritage 2011. "English Heritage's MIDAS standards page." Accessed June 30. <http://www.english-heritage.org.uk/professional/archives-and-collections/nmr/heritage-data/midas-heritage/>

Internet Archaeology 2005. *Internet Archaeology* 18. Accessed June 30. <http://intarch.ac.uk/journal/issue18/>

Jeffrey, S., Richards, J., Ciravegna, F., Waller, S., Chapman, S., and Zhang, Z. 2009. "The Archaeotools project: faceted classification and natural language processing in an archaeological context." *Philosophical Transactions of the Royal Society* 367:2507–2519.

Kenny, J., and Richards, J. D., with a contribution from Waller, S. 2005. "Pathways to a Shared European Information Infrastructure for Cultural Heritage." *Internet Archaeology* 18. http://intarch.ac.uk/journal/issue18/kenny_index.html.

Kenny, J., Kilbride, W. G., and Richards, J. D. 2003. "Enter the ARENA: preservation and access for Europe's archaeological archives" In *Computer Applications and Quantitative Methods in Archaeology 2002*, edited by M. Doerr, and A. Sarris, 349-353. Archive of Monuments and Publications, Hellenic Ministry of Culture.

McManamon, F. P., and Kintigh, K. W. 2010. "Digital Antiquity: Transforming Archaeological Data into Knowledge." *The SAA Archaeological Record* 10 (2):37-40.

MOLAS 2011. "Museum of London, Spitalfields case study site." Accessed June 30. <http://www.museumoflondonarchaeology.org.uk/News/spitalfieldscasestudy.htm>

[museumoflondonarchaeology.org.uk/News/spitalfieldscasestudy.htm](http://www.museumoflondonarchaeology.org.uk/News/spitalfieldscasestudy.htm)

Richards, J. D. 2008. "Managing Digital Preservation and Access: The Archaeology Data Service." In *Managing Archaeological Resources: Global Context, National Programs, Local Actions*, edited by F. P. McManamon, A. Stout, and J. A. Barnes. Walnut Creek: Left Coast Press.

Spielmann, K. A., and Kintigh, K. W. 2011. "The Digital Archaeological Record: The Potentials of Archaeozoological Data Integration through tDAR." *The SAA Archaeological Record* 11 (1):23-25.

TAG 2011. "TAG Project Page." Accessed June 30. <http://archaeologydataservice.ac.uk/research/tag>

TAG Search Interface 2011. "TAG Search Interface." Accessed June 30. <http://archaeologydataservice.ac.uk/TAG/intro.jsf>

tDAR 2011. "tDAR." Accessed June 30. <http://www.tdar.org/>

Waller, S. 2005. "Future Connections: The potential of Web service and Portal technologies for the historic environment." *Internet Archaeology* 18. Accessed June 30. http://intarch.ac.uk/journal/issue18/waller_index.html.