Archeologia e Calcolatori
15, 2004, 129-144

THE DANUBE IN PREHISTORY IN THE DIGITAL AGE: TOWARDS A COMMON INFORMATION ENVIRONMENT FOR EUROPEAN ARCHAEOLOGY

1. Introduction

The boundaries of the nation state are a recent innovation. They were unknown to our prehistoric ancestors. The study and management of archaeology should be well placed to benefit from increased co-ordination within and around the European Union: but the impact of European harmonisation upon archaeology has been surprisingly modest. Nowhere is this more obvious than in the supply and exchange of heritage information. Our archaeology may be continuous, but the dominion of those agencies charged with protecting and recording it are precisely curtailed. Information, the principal tool of responsible management and conscientious research, is fragmented among miscellaneous authorities.

Consider V.G. Childe's masterly survey of the Danube in Prehistory (Childe 1929). Written at a time when travel and communication were arduous and access to information a rarity, it presented a comprehensive synthesis that combined new ideas with the details of latest research. What would happen if Childe were to attempt this task again using modern communications? Though the amount of relevant information is infinitely greater than in the 1920's, modern communications technologies – especially the Internet – should make this colossal task easier. In reality, the din of competing websites, electronic publications, discussion lists, databases and GIS services would more likely hinder than assist his research.

Better access to information could improve heritage management right across Europe, and be an impetus to better research and learning. But uniting and taming such a diverse set of information tools is not a task to be taken on idly. This paper reviews continuing innovations in the electronic dissemination of archaeological information. It will be seen that technical interoperability can be achieved with relative ease, but that technology alone cannot provide a solution to organisational and semantic diversity. Archaeology is coherent at a European scale: how can we make our information resources coherent at that scale too? A number of possible solutions and the work required to achieve them are outlined. At the end it will be noted that the biggest challenge is not technical but organisational.

2. Context: towards a boundary-less Europe

The European Union has transformed the institutions of political and economic life in the continent. This move has had its most obvious impact in archaeology through the activities of the European Association of Archaeologists. The EAA is one of many organisations that remind us of the multi-national relevance of archaeology, but it is supported by a wider political and cultural infrastructure than many (e.g. Willems 2000). European archaeologists are converging in policy and practice: shared endeavours that extend beyond bonhomie towards harmonisation. Projects and declarations, like Planarch (Clarke 2002), the Esslingen Code (Croft, Thomas 2001) and the Sevilla Declaration (García Sanjuán, Wheatley 2002) have brought together local heritage managers to share expertise and develop co-operative, complimentary heritage management policy.

This emergent harmonisation of practice and policy is unlikely to threaten the institutional frameworks in which our research is executed and our common heritage is managed. So we cannot expect, nor should we seek, a rapid integration of information through organisational merger. Practical solutions to the fragmentation of information in archaeology are more likely to be found in the world of information science.

It could be argued that the problem of fragmentation is a ruse to justify superfluous research. In 1992, Henrik Jarl Hansen speculated on the prospect of creating a unified index to European archaeology, and identified benefits that could follow from it (Hansen 1993). To some extent, the Internet provides the sort of network that Hansen envisaged.

Yet all is not as it seems. Firstly, only a few heritage managers are able to provide online access to detailed heritage management records. There are technical and legal impediments to overcome as well as financial ones. Moreover, conventional search engines are poorly equipped to answer the detailed questions that we ask routinely. They lack authority; they lack currency; and they are not easy to track. Internet searches frequently produce records that may be irrelevant or lacking the high standards of academic rigour that professional research requires. Internet search engines tend to favour older records over newer ones: recent additions or new material may simply not be indexed at all, while old information, which may be out of date, will be ranked more highly. Finally, general Internet searches produce records that are hard to track. It is a common frustration that websites disappear, change or are renamed through time. These problems are more than just minor inconveniences. They could make the difference between good management decisions and appalling ones: between valid and invalid research conclusions. The Internet, as currently structured, does not resolve the problem of fragmentation.

The prospects for the future remain bright: solutions are available, but perhaps more importantly, the will and the opportunity both exist. Digital resources generally, and cultural heritage data sets in particular, have captured

the attention of the European Union. Three distinct but connected policy areas mean that the electronic deployment of cultural heritage information could see significant investment in the coming years: the development of a single European Research Area (EC 2000); the development of a Europe-wide "e-government" agenda (EC 2003) and the harmonisation of cultural heritage digitisation policies (National Representative Group 2003). This last area – encapsulated in the Lund Principles – deserves particular attention.

The most recent action plan for cultural heritage in Europe calls for the creation of a co-ordination mechanism for digitisation programmes across member states. This follows a meeting in Lund in Sweden where an expert panel drawn from each member state adopted a set of basic principles to underpin future strategy for the creation and release of heritage information. These principles commit member states to the establishment of national fora to co-ordinate local activities, and to the dissemination and adoption of good practice with digital resources through identified centres of competence. The principles seek better bench-marking of digital products with identifiable quality indicators and enhancements to access mechanisms. Perhaps most importantly, the Lund principles call for member states to: «…optimise the value and develop shared visions of European content, by developing criteria and a framework for an EU collaboration plan… The plan should aim at establishing an eCulture infrastructure for access to digitised cultural and scientific heritage through identifying added value conditions for European content» (National Representative Group 2003, 199).

The implications for archaeology are three-fold. Firstly, if we wish to secure future funding to continue the release of cultural heritage information sets, our representations to do so will be looked upon more favourably where there is co-ordination. Secondly, those engaged in computing will be expected to work to formal standards. Finally, and perhaps most importantly, we should be preparing ourselves for participation in a new, wider infrastructure of cultural content. That infrastructure may not exist yet, and may indeed be several years away; but the greatest impact will be had by those data collections that are co-ordinated at the outset. Of all scientific and cultural endeavours, archaeology should be well placed to lead this development. Diverse in content but continuous in its distribution, the "added value conditions" of archaeological data are surely among the most resilient.

3. Distributed computing: technology to the rescue?

These high-level policy ambitions are long-term goals, but are built on short-term actions. The prospects are promising and perhaps more importantly much of the fundamental research has already been done. Exchange protocols and metadata standards have received a lot of attention in the last ten years. There is also widespread agreement on XML as a transport and storage format, though the examples of use remain few.

3.1 *Protocols for distributed computing*

Agreement on the formal machine-to-machine mechanisms by which information is exchanged is essential. Work by the ADS and partners on two open-source protocols prove their promise for the future: the Z39.50 protocol; the OAI protocol.

The Z39.50 protocol was developed to allow librarians to present multiple and distributed library catalogues through a single interface (Miller 1999). A number of consortia have been formed to provide just such a service to a number of significant data collections for higher education. By adopting a common profile and a common protocol, libraries and archives make their collections directly and easily available to each other via cross searching. That way a large number of library and archive catalogues can be interrogated simultaneously. Library catalogues provide an obvious example of how diverse and distributed databases can be searched simultaneously. Museum catalogues provide another example. Thus, a profile for the exchange of museum information – the CIMI profile – extends the functionality of Z39.50 searching to include the sorts of locational and temporal information commonly found in archaeology.

This protocol has also been used in the context of heritage management information. Launched in January 2002, HEIRPORT (http://ads.ahds.ac.uk/heirport/) cross searches 4 distributed databases in 4 different parts of the UK. It searches the National Monuments Record of Scotland (CANMORE), the Scottish Cultural Resources Access Network (SCRAN), the Portable Antiquities Scheme (PAS) and the main catalogue of the Archaeology Data Service (Austin *et al*. 2002a and b). Each of these databases has its own web presence: the short records presented via HEIRPORT link back to the complete originating record. Thus, rather than competing, HEIRPORT provides a sophisticated presentation layer through which users can access and compare data from numerous sources. Each of the "targets" can in turn be queried by any other Z39.50 server, so that the same data can be presented in different contexts. Information is transmitted in XML and transformed using style sheets before being displayed. This normally means that the results are presented for display in a standard web browser in HTML, but in tests, results were also displayed in WML for display on WAP devices.

The Z39.50 protocol has the advantage that once established, data does not need to be exchanged: indeed the server itself holds no information of its own, only the names and the configuration of the clients it seeks to query. Rather than swapping data, the participants need only swap and configure processes. Moreover, because the systems remain independent, targets may be accessed by more than one server. The fact of mutual agreement imposes a degree of quality control, thus overcoming many of the authority issues

associated with conventional Internet search engines. HEIRPORT thus gives a taste of what is possible where heritage managers present information collaboratively. It is a precursor of what a common information environment in archaeology will look like (Kilbride 2003).

Although Z39.50 is a flexible and powerful tool, there is no doubt that it is technically complicated and processor intensive. Its greatest strength – the live feed of current information – can also become a weakness when network services are congested or individual servers unavailable. An alternative solution to the same problem of cross searching distributed databases is offered in the Open Archives Initiative Protocol for Metadata Harvesting (OAI hereafter). This recently released protocol offers very similar functionality to Z39.50, but is generally held to be less demanding, and certainly makes more efficient use of network time. OAI was developed to support the development of E-print repositories of pre-print articles and thesis, which have had a remarkable impact on scholarship in the sciences (Van de Sompel 2000).

The OAI protocol has many similarities to Z39.50. A central server allows simultaneous searching through a pre-defined metadata set which in turn derive from remote targets. Again there is a concordance between the fields in the remote target, and the shared metadata standard. However, in OAI, the server harvests metadata, parses it and appends the results of its harvesting to a single unified database. Thus, rather than broadcasting searches as and when a user requires, the OAI server responds to user requests by searching its own local metadatabase. The frequency of harvests is not defined, so can be agreed to suit the frequency of changes in the originating database. As with Z39.50 the metadata gathered links back to the originating record or digital object, and any given OAI enabled database can be searched by any other OAI server, provided they agree on a metadata schema. Like Z39.50, OAI extends web databases by providing a new presentation layer to different users, and in the context of other databases.

OAI has a great deal to commend it in terms of cultural heritage computing. For example, because the metadata is copied to a central location and searched by the same system, the amount of network traffic it generates is controlled. Response times to user requests can be more rapid than an equivalent Z39.50 search, and are more robust. These advantages are offset by problems of currency: metadata presented this way is only as current as the last harvest. Thus, establishing an OAI server locks developers into a continuous cycle of harvesting to ensure that both sets of data are synchronised.

More problematic for OAI are issues to do with metadata schemas. Z39.50 has a mature and fully developed set of profiles to support cross searching. OAI can support any number of schemas, but defaults to the use of unqualified Dublin Core. Dublin Core is a useful standard, but unqualified it can be problematic across more than one data set.

Other machine-to-machine protocols for information exchange supplement and extend cross searching activities. For example, the Resource Description Framework (RDF) news syndication service can be used to feed multiple websites simultaneously (Miller 2003). The Simple Object Access Protocol (SOAP) is increasingly used to extend Z39.50 services. SOAP is part of the Web Services Technology (WST) framework designed to reduce the cost and increase the modularity of application integration. Other emerging elements of this suite of products includes Universal Data Description and Integration standard (UDDI) and the Web Services Definition Language (WSDL). The Web Services Technology standards are designed to provide "wrapper" applications to allow remote systems to interact directly, thus reducing development costs. Though of a very different origin from the library and archive protocols of Z39.50 and OAI, it has been argued that there is conformance in the functions that are provided (Powell, Lyon 2002). Though not yet implemented in any archaeological contexts at this time, the continued development of Web services technologies, allied to other open standards hold considerable promise for the future.

In conclusion, there is already an embarrassment of riches as far as protocols are concerned for the interoperation of distributed information systems. Some are already mature: others will mature rapidly and will be adopted by the archaeological community. We have a very diverse and powerful set of tools at our disposal.

3.2 *What is to be exchanged: metadata standards*

Metadata standards underpin most cross-searching activities (Miller, Wise 1997). ArchSearch, the main catalogue of the Archaeology Data Service, shows that complicated protocols are not necessary when appropriate metadata standards are adopted (http://ads.ahds.ac.uk/catalogue/). By mapping the information held in diverse records, and then concatenating them into a single index, ArchSearch is able to present a single online interface to a significant number of the UK's heritage management databases. Based on an implementation of the Dublin Core metadata element set (Miller, Greenstein 1997), it provides "item level" records for sites and monuments that link to specific site archives where they exist.

Recently, funding from English Heritage has allowed the ADS to look at information flows in English archaeology, and develop systems that can cut out much of the duplication of effort currently present (Hardman, Richards 2003). The OASIS project provides a theoretically simple solution to problems of duplication in the entry and supply of heritage management information in the UK. It provides a mechanism whereby field units, who are the creators and originators of much of the data can for the first time supply their findings to local government in electronic form, with a simple mechanism for supply those records to any other interested agency.

OASIS touches every part of the information flow from initial fieldwork to publication – recognising the need for different types of validation, the complexity of intellectual property rights, and the needs of different management functions in archaeology. Amongst the results of this work is a trial of a single, unified mechanism for exchanging information between different agencies, based on electronic supply. An agreement on what fields are needed and how they should be presented has provided what is in essence an extended metadata set of 45 elements, with an online tool for units and managers to create, edit, amend, download and distribute metadata. This sets the foundation for an XML schema for archaeological information and may yet prove to be more widely applicable than just for heritage management in England.

This is just one of many recent developments that are contributing to the development metadata standards. Indeed, it is the plethora, not the dearth of metadata schema that may prove a challenge in the future. The proliferation of metadata standards presents an opportunity to develop a higher level ontology in which varied concepts can be expressed, and a semantic web where automated tasks are done by agents on our behalf (Steemson 2003). The best example of an ontology for archaeology is the CIDOC Conceptual Reference Model (CRM) which allows different standards to be reconciled (http://cidoc.ics.forth.gr/). This offers an ontology of 81 classes and 130 properties which describe the concepts and relationships between concepts current in heritage management. This ISO standard allows diverse datasets to be interoperable with any other by providing a basic, universal set of concepts that are represented by all information systems. In doing so, it provides an alternative approach to the cross searching problem identified at the outset. The CRM shows the need and potential for cross-walking technologies based on a sound and systematic ontology.

3.3 *XML in heritage information*

Archaeology has been relatively slow to see the advantages of XML, and there are only a few examples of its use for peculiarly archaeological purposes. XML underpins Z39.50, RDF and OAI protocols, and so has been used widely in Heirport. It is one of the transmission formats for records generated from the OASIS data tools. There is every reason to explore the use of XML to mark up and present the fundamental archaeological records too. If an excavation report were presented in an archaeological XML format, then it would be possible for field units to use a single XML file to generate all manner of information. An appropriate series of XSLT style sheets could then represent the parts of the report required by the individual user: index records for heritage managers, summary reports for general readers; structural reports for researchers; finds reports for specialists and so on.

Text markup of the type described here is advanced in other fields, including linguistics, mathematics, physics and music. Tools like the Text Encoding Initiative (TEI) Schema or Encoded Archival Description schema provide frameworks which archaeology could adapt (Morrison *et al*. 2000). A combination of the OASIS metadata schema for the high level archaeological information with the hierarchical approach of the text encoding initiative for marking up the body of a text could go some way to providing a complete information parcel in archaeology where data and metadata need not be separated. The graphic and geographic nature of much archaeology may present challenges to TEI approach, though tools like Geospatial Markup Language (GML) and Scalable Vector Graphics (SVG) are likely to provide support in these fields.

There is widespread support for the development and use of XML in archaeology, but too little active development work (though see Meckseper 2001; Sugimoto 2002). It is clear that many of the basic tools are present, in the form of OASIS records, TEI and other schema. These tools need to be more widely known and evaluated in archaeology, and will have to be coordinated if the benefits of XML are to be exploited more fully.

4. Why technology (alone) is not the answer

The adoption of various configuration of protocols and metadata standards are clearly welcome, but it would be a mistake to think that all we need do is add data and hope that this will allow the sort of boundary-less research envisaged at the start. Protocols and standards may allow us bring information together, but they won't make it coherent. On the contrary, a recent progress report on the Lund Principles points clearly to issues that need to be addressed: «Even if much work has already been done on metadata and standards, the next big challenges are technical and organisational interoperability and multilinguality» (National Representative Group 2003, IX).

4.1 *Language problems: home and away*

Language problems start at home. Archaeology is poorly served by universal library subject headings that tend to be too generalised to offer internal classifications. At a very detailed level, there are a variety of standard terminologies which can be used to unite inventories. The more sophisticated of these tools provide hierarchical relationships between terms creating groups of related terms (e.g. Rchme, English Heritage 1998). They are aimed in particular at specific vocabulary controls and normally relate to discrete phenomena. These tools are vital for the concordance of diverse data sets, but they also provide powerful browsing tools for information retrieval.

A number of initiatives are currently exploring multilingual thesauri, but progress is necessarily painstaking. Specialist projects like HEREIN have been working to provide multilingual thesaurus to a help access a range of heritage management policies, while a multilingual thesaurus of the European Bronze Age deals broadly with the details of a narrow topic.

The Archaeological Records of Europe Network Access (ARENA) project provides some insight into the problem of making terms mutually comprehensible. One aim of this path finding project is to pilot a multilingual portal for heritage information (Kenny, Kilbride, Richards 2003). In order to provide useful search mechanisms, the partners are mapping items in a number of monument inventories to the 18 highest-level terms in the Thesaurus of Monument Types (Rchme, English Heritage 1998). Partners have translated the 18 highest level terms from the thesaurus – agriculture, civil, commemorative, communications, defence, domestic, education, gardens and parks, health and welfare, industrial, maritime, recreational religious, transport, and water systems – into Polish, Norwegian, Danish, Icelandic, Romanian. These terms are then applied to partners' datasets, allowing a researcher to find all the monuments across these six different countries that match the highest-level term.

This work shows both the benefits and the drawbacks. On one hand a whole range of data can now be compared with a degree of confidence: on the other the breadth of the categories makes that comparison problematic. The more detailed the terms become, the more difficult it will be to translate the concepts that underpin them.

4.2 *Time and space as classification*

Part of the problem of multi-lingual computing for archaeology is the lack of any external industry or innovation which will solve the problem for us. However, language is not the only means to classify archaeological information, nor is it even always the most useful (Kilbride 2002b). Archaeological information is pre-discursively geographic: librarians and information scientists are only slowly coming round to the idea of mapping as a form of classification. In addition archaeological data has a strong temporal component. Though more treacherous than it first appears, this temporal component also supports classification.

This is another area of research by the ARENA partnership. Each of the ARENA partners has agreed to release site-based information to a central portal through either the Z39.50 or OAI protocols. Cross searching these records – even with the eighteen agreed high level terms – is problematic as the records are very numerous. However, the metadata scheme used by the six partners includes a numeric map reference for each site, either expressed in decimal latitude and longitude, or dynamically exported into latitude and

longitude from the local co-ordinates system. Thus, with the inclusion of detailed mapping, it becomes possible to define searches that are more useful than the simple eighteen high level terms.

This task is more convoluted than it may first appear. There is bound to be a degree of imprecision introduced by the conversion to latitude and longitude which will be particularly noted at the convergence of different mapping systems. The problem is not whether there will be imprecision: it is whether such imprecision is within tolerable limits. Similarly, the accuracy of the native co-ordinate information cannot be taken for granted: the degree of acceptable error is not yet known, and whether researchers in different counties share a similar tolerance of positional accuracy can only be hoped. This raises a supplementary question on the need (or not) for confidence-metadata. Finally, ARENA is only experimenting with simple point data. It is to be hoped that a similar set of tools could be built out of linear and polygonal information – but this in turn will require investigation into the construction of spatial databases. The ARENA project is only intended as a way-finding initiative, and the very fact that these subsequent refinements are discussed here implies useful progress on more fundamental issues.

Less well exploited by archaeologists, and intellectually more complicated are the temporal characteristics of heritage information. This is another aspect of the work of the ARENA partnership. The temporal characteristics of each inventory has been analysed and each of the terms used have been compared to a numerical timeline. Thus, it is to build a query to retrieve information from six different agencies about archaeological sites that belong to the same century or millennium. This research has presented a number of issues, some of which were anticipated and others that were not obvious at the outset.

Period terms may seem similar across archaeology: but as the area increases in size, so the utility of period descriptions declines. Thus, while the term "Iron Age" may mean a specific period in the Highlands of Scotland, it refers to a radically different period in the south of England. Archaeological periods are closely related to geography, and the conclusion is that any system which uses periods to classify archaeological data should be cognisant of geographic context too.

The temporal characteristics of a period or site are the subject of debate. ARENA partners have agreed to present all results that might be interpreted as one or other period, and ask the user to evaluate them. Perhaps more problematic is how to deal with antiquity in an ancient landscape. Although sites can be multi-period, there is no weighting between the periods. This becomes an issue because the very fact of survival means that many sites can reasonably be said to belong to every period *since* construction. So a complex archaeological site like the ramparts of Iron Age hill fort at Danebury

in Wiltshire may reasonably be described as part of the Roman, Anglo-Saxon, Medieval, Post-Medieval and Modern landscapes. Yet, if one were to ask the portal about the medieval landscape of Southern England, Danebury, which must have been part of that landscape, is not retrieved.

Dating archaeological sites is an expert task and almost every sort of evidence is used, including documentary study, site morphology, artefact typology, radiocarbon dating, dendrochronology, toponymics and studies of vegetation pattern. The confidence associated with chronology can vary tremendously. Some sites can present radiocarbon dates expressing confidence as a statistical probability; others rely on the associative dating of a type-site or artefact assemblage: others on immediate stratigraphic relationships with dated material. There is no common currency by which confidence in dating is reported in archaeology, and until there is there are no mechanisms by which users can be guided.

4.3 *Levels of description: collections as items*

The focus of this article, and much of the work being carried out, has been on the immediate gratification of requests for research data. Paradoxically by seeking to develop tools that will ultimately provide instant answers, we may be setting ourselves an arduously long and risky development path. By stepping back from the problem slightly, we may find that this is an unsatisfactory approach. It is ultimately based on the dubious premise that European archaeology actually knows where to look for its best data assets. It is likely in fact that very few European archaeologists are aware of many of the data sets beyond their own specific area of expertise, and have no obvious mechanism of remedying that ignorance.

Returning to Childe's *Danube in Prehistory* (Childe 1929) we should realise that the Danube flows through (or borders) Germany, Austria, Slovakia, Hungary, Croatia, Serbia, and Bulgaria before emptying into the Black Sea in its massive Romanian delta that reaches north to the border with Ukraine. Without even considering watersheds and tributaries, we are already required to negotiate access to information from nine different national governments and a myriad of local, regional and provincial authorities. It is a significant challenge even to know which organisations we should contact, and where the important datasets are curated. Perhaps as well as trying to combine monument inventories, we should at the very least establish an inventory of inventories.

There are two examples of collection level metadata catalogues in archaeology in the UK, and their use delineates the prospective uses of collection description in archaeology more generally: the ARCHway suite of tools and the HEIRNET register. ARCHway is a union catalogue of archaeology journals held by UK university libraries (http://ads.ahds.ac.uk/catalogue/ARCHway.html). Partner libraries started by compiling lists of their archae-

ology journals noting which volumes of each title was held. These records were combined to create a master index of the 2500 archaeology journals available to researchers in the UK, along with details of which libraries hold those journals, and which years are available. The result is that one search can tell librarians and researchers alike where to find archaeology journals, and where the significant overlaps of holdings are.

In addition, the tables of contents of 13 journals were then digitised so that every single article and note in every issue of the most important journals now has its own record. Thus, though it only works for a small proportion of the total, it is possible to start with a generalised "collection-level search" through the whole of the archaeology journal holdings in UK universities, to details of where to get hold of each journal, then move to an "item-level record" for each article in that journal giving page numbers, title and author. In some cases it provides those articles in digital form.

The ARCHway experience is exemplary because everyone who has studied archaeology for any length of time has had the problem of trying to track down an obscure article or journal that is not immediately to hand. It is also telling that researchers find it frustrating that the system only occasionally gives them immediate access to the full text of the articles in question.

The implication of collection level description for heritage information is broader though. If each of the national, provincial and regional agencies were to describe its information holdings, with a record of their roles, responsibilities and contact details then the opportunities for research would be greatly expanded. If that record also described the sorts of data maintained by these agencies and which parts are available to researchers, then both researchers and heritage managers would benefit. Researchers would find it easier to obtain the information they need, presenting appropriate and sustainable questions of the appropriate authorities. They would thus be able to progress more rapidly to the more complex and rewarding work of researching the archaeology itself rather than the complex organisational structures through which it is managed. Managers would be given a network of fellow professionals to complement and improve their own work, which could provide a basis for sharing experience and information. One can only presume that such co-ordination would also benefit the heritage concerned (Kilbride 2002a).

An example of such a collection-description framework in the context of heritage management information exists in the HEIRNET Register (Chitty *et al*. 2000 and http://ads.ahds.ac.uk/heirnet/index.cfm). This provides an index to all the available heritage information sets in the United Kingdom, and details of the organisations that maintain them. This is not restricted to online information sets: indeed the vast majority of the information described is offline, and only available to those with appropriate credentials, or willing to make the journey to consult them. This has three important consequences.

Firstly, the heritage managers that maintain the records retain a large degree of discretion as to who accesses the records and which parts they are allowed to see. Secondly the records produced are in fact very flexible and can be used for very many alternative purposes. Thirdly, and perhaps most importantly, it places online and offline records on equal footing, and is thus relatively inexpensive in both political and financial terms. Those agencies unable or reluctant to release full records online can still participate. It thus enables research without deflecting heritage managers from their core responsibilities. In short, the sorts of collection description records available through the HEIRNET register provide a strong support for researchers and heritage managers alike, at relatively little cost.

This article has focussed on the integration of monument inventories at an item level. This is a significant undertaking which may take many years to complete: a simpler approach of indexing the inventories that already exist would be less difficult to achieve, though similar problems would inevitably arise. Language problems for cross searching would still be present, and the problems associated with geo-spatial and temporal classification would still need to be resolved. A greater risk to such an index may however come from users who may be deterred by the fact that specific information which they seek is not immediately accessible online, only an email address or telephone number through which they may obtain more information. As ARCHway shows, users generally want more than just index records. ARCHway also shows, however, that without these index records, the item level objects may make no sense. Collection level records could be useful in the short term and will be needed in the longer term.

5. Conclusions: the Danube in Prehistory in the future?

This article has highlighted the tools available to support more thoroughly integrated research and management of Europe's archaeology. In doing this, it describes a vital and sophisticated research community that can provide versatile solutions for a variety of organisational needs. The future holds more as the semantic web, the Grid and other network based technologies continue to revolutionise how we present, compare, integrate and analyse information. Perhaps a more important and difficult question is what in fact we need and want to do with these technologies.

While we may have useful tools in the form of protocols and agreed metadata standards, there is still a lot of work that will be required on related issues. Language problems cannot be overcome by brute force of technology, while issues of what level of description is practical hint at a broader discussion of what sort of information environment we really want for archaeology.

Even at this level however, our concentration has been on the technical and informational, not the organisational or political. While there is some degree of consensus that there is much to be gained, and while there may well be an opportunity in terms of funding, it is clear that there is need for a wider debate and discussion and agreement among European archaeologists before successful progress can be achieved. Any technology-led project that fails to take account of its core user community is ultimately doomed to fail. Consultation of users improves the core system, but it also creates a sense of ownership over the product. Both of these virtues are required if such a system is to succeed. That in turn raises any number of questions about how we build and sustain momentum behind such a project, how we involve the large number of stakeholders. These invisible obstacles may yet be larger than the technical ones upon which we tend to focus.

William Kilbride
Archaeology Data Service
Department of Archaeology
University of York

REFERENCES

Austin T., Pinto F., Richards J., Ryan N. 2002a, *An interoperable portal for the historic environment*, in E. Neuhold, L. Kalinichenko (eds.), *Proceedings of the Third DELOS Workshop: Interoperability and Mediation in Heterogenous Digital Libraries*, DELOS Network of Excellence on Digital Libraries Workshop Series, Darmstadt (http://www.cs.ukc.ac.uk/pubs/2001/1262/index.html).

Austin T., Pinto F., Richards J., Ryan N. 2002b, *Joined up writing: an Internet portal for research into the historic environment*, in G. Burenhult, J. Arvidsson (eds.), *Archaeological Informatics: Pushing the Envelope. Computer Applications and Quantitative Methods in Archaeology*, *CAA 2001*, Oxford, Archaeopress, 243-252.

Childe V.G. 1929, *The Danube in Prehistory*, Oxford, Clarendon.

Chitty G., Baker D., Richards J.D., Robinson D.J. 2000, *Mapping Information Resources: A Report for HEIRNET*, Council for British Archaeology (also available online at: http://www.britarch.ac.uk/HEIRNET/rep1.html; last visited 19/12/03).

Clarke C.P. (ed.) 2002, *Protecting the Past for the Future: the Development of SMR's in the Planarch Project Region and Beyond*, European Regional Development Fund, Chelmsford.

Croft B., Thomas R. 2001, *"The Esslingen Code" – a European code for urban archaeology*, «The European Archaeologist», 16, 14.

EC 2000, *Towards a European Research Area*, *Communication from the Commission to the Council, the Parliament, the Economic and Social Committee and the Committee of the Regions*, *Commission of the European Communities*, Brussels (ftp://ftp.cordis.lu/pub/documents_r5/natdir0000001/s_1372005_20010125_143514_C001190en.pdf; last visited 15 December 2003).

EC 2003, *The Role of E-government for Europe's Future*, *Communication from the Commission to the Council, the Parliament, the Economic and Social Committee and the Committee of the Regions*, *Commission of the European Communities*, Brus-

sels (http://europa.eu.int/information_society/eeurope/2005/doc/all_about/egov_communication_en.pdf; last visited 15 December 2003).

García Sanjuán L., Wheatley D. (eds.) 2002, *Mapping the Future of the Past: Managing the Spatial Dimension of the European Archaeological Resource (Sevilla 1999)*, Sevilla, Universities of Sevilla and Southampton and Andalusian Institute of the Historical Heritage.

Hansen H.J. 1993, *European archaeological databases: problems and prospects*, in J. Andresen, T. Madsen, I. Scollar (eds.), *Computing the Past: Computer Applications and Quantitative Methods in Archaeology 1992*, Aarhus, Aarhus University Press, 229-237.

Hardman C.S., Richards J.D. 2003, *OASIS: dealing with the digital revolution*, in M. Doerr, A. Sarris (eds.), *The Digital Heritage of Archaeology: Computer Applications and Quantitative Methods in Archaeology 2002*, Heraklion, Hellenic Ministry of Culture, 325-329.

Kenny J., Kilbride W.G., Richards J.D. 2003, *Enter the ARENA: preservation and access for Europe's archaeological archives*, in M. Doerr, A. Sarris (eds.), *The Digital Heritage of Archaeology: Computer Applications and Quantitative Methods in Archaeology 2002*, Heraklion, Hellenic Ministry of Culture, 349-354.

Kilbride W.G. 2002a, *Information and the management of Europe's historic environment*, in Clarke 2002, 69-76.

Kilbride W.G. 2002b, *Mapping the heritage information landscape: interoperability and geo-spatial description in the management of Europe's historic environment*, in García Sanjuán, Wheatley 2002, 115-127.

Kilbride W.G. 2003, *One for all and all for one: towards a common information environment*, «Archaeology Data Service News», 14, 6-7.

Meckseper C. 2001, *XML and the Publication of Archaeological Field Reports*, unpublished MSc Thesis, University of Sheffield.

Miller P. 1999, *Z39.50 for all*, «Ariadne», 21 (http://www.ariadne.ac.uk/issue21/z3950/; last visited 19/12/03).

Miller P. 2003, *Syndicated content: it's more than just some file formats*, «Ariadne», 35 (http://www.ariadne.ac.uk/issue35/miller/; last visited 19/12/03).

Miller P., Greenstein D. (eds.) 1997, *Discovering Online Resources across the Humanities: A Practical Implementation of the Dublin Core*, UKOLN, Bath (also available online at: http://ahds.ac.uk/public/metadata/discovery.html; last visited 19/12/03).

Miller P., Wise A. 1997, *Why metadata matters in archaeology*, «Internet Archaeology», 2 (http://intarch.ac.uk/journal/issue2/wise_index.html; last visited 19/12/03).

Morrison A., Popham M., Wikander K. 2000, *Creating and Documenting Electronic Texts, Oxford Text Archive*, AHDS Guides to Good Practice, Oxford, AHDS (also available online at: http://ota.ahds.ac.uk/documents/creating/; last visited 19/12/03).

National Representative Group 2003, *Co-ordinating Digitisation in Europe: Progress Report of the National Representative Group Co-Ordination Mechanisms for Digitisation Policies and Programmes 2002*, Rome, European Commission.

Powell A., Lyon L. 2002, *The JISC information environment and web services*, «Ariadne», 31 (http://www.ariadne.ac.uk/issue31/information-environments/; last visited 19/12/03).

Rchme, English Heritage 1998, *Thesaurus of Monument Types: a Standard for Use in Archaeological and Architectural Recording* (second edition), Swindon, Royal Commission on the Historic Monuments of England (also available online at: http://www.english-heritage.org.uk/thesaurus/thes_splash.htm; last visited 19/12/03).

Steemson M. 2003, *Digicult's Expert 13 tangle with the Semantic Web*, in *Towards a Semantic Web for Heritage Resources*, «Digicult», 3, 14-24.

SUGIMOTO G. 2002, *XML Application for Archaeological Image Database: an Example of Herefordshire SMR Aerial Photographs*, unpublished MSc Thesis University of York.

VAN DE SOMPEL E. 2000, *The Santa Fe Convention of the open archives initiative*, «D-lib Magazine», 6 (http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html; last visited 15/12/2003).

WILLEMS W.J.H. 2000, *Challenges for European Archaeology: Report of the Inaugural Meeting of the Europae Archaeologiae Consilium at the Council of Europe in Strasbourg (Strasbourg 1999)*, OcenW.

ABSTRACT

Europe is a very old and very small continent. The accession of ten new states to the EU in spring 2004 reminds us that the political boundaries we police and survey would have made little sense to the ancient populations who moved freely across our frontiers. Our disparate national and local heritage services represent different traditions and experiences of researching, recording, presenting and managing what should be among our principal assets. This diversity risks undermines research and conservation, it inhibits international strategies for heritage management and institutionalises anachronism. Can information technologies support the EU's stated aim of creating a single European research area for archaeology? This paper investigates the long-standing question of how different archaeological data sets in different parts of Europe may be aligned more closely to support research learning and teaching. It identifies emerging technologies to for resource discovery, integration and delivery, placing these in the context of organisational evolution. It asks how these organisations and technologies might work together to support archaeological information at a continental scale.