

Indexing and Retrieving Archaeological Resources on the Internet - A prototype Multilingual Thesaurus Application

Martijn van Leusen

Groningen Institute of Archaeology
Poststraat 6 9712 ER Groningen, The Netherlands
e-mail: p.m.van.leusen@let.rug.nl

Abstract

This paper introduces the concept of using a thesaurus for structured indexing and retrieval of archaeological information on the Internet, defines parameters for the implementation of a multilingual web-based thesaurus application, and describes the prototype under development by the ArchTerra Consortium.

Keywords: thesaurus, Internet

1. Introduction

The field of Archaeology shares in the general exponential growth of the amount of information that is being published on the Internet. Whereas the first such publications in Europe date to “gopher” sites of the early 1990s, the major search engines currently report anything up to 100,000 pages relevant to European archaeology. Over the intervening years many have lamented the resulting chaos and lack of quality control, although knowledgeable users could still reduce the chaos by using the various information filtering mechanisms offered by web index sites such as AltaVista. Champion (1997) and Van Leusen et al. (1996) provide a good overview of the situation in 1995/6, and over the past five years the general public too has been educated in the use of such mechanisms.

As the Internet became a potential source of information for the professional archaeologist too, the need for more sophisticated retrieval mechanisms (and their obverse, indexing mechanisms) becomes ever more pressing. Among the early solutions were Internet “guides” such as ARGE,¹ which provide access to a manually indexed subset of web resources, and metadata schemes such as that of the Dublin Core,² which allow the author of a web page to include descriptive information which can then be retrieved by web indexing robots. While these solutions provide short-term relief, the need for long-term, cross-national, and effective solutions to the problem of finding information on the Internet has been well recognised. The AQUARELLE project “Sharing cultural heritage through multimedia technologies”³ represents an industry-driven approach to the problem, while professional archaeologists themselves have opted for the distributed, co-operative approach represented by the ArchTerra Consortium.

The cultural and “information society” programs sponsored by the European Commission as part of the 4th and, recently, 5th frameworks provide important means for experimenting with the development of solutions. Straddling the two frameworks is the ArchTerra project⁴ (1999-2000), which aims to extend the invisible web of European archaeology (which is, after all, a largely western European and English-language web) to eastern Europe. In addition to putting eastern European archaeology online, the ArchTerra consortium develops distributed and multilingual tools to facilitate the publication and retrieval of archaeological infor-

mation on the Internet. One of these is the online multilingual thesaurus system of which a prototype is presented in this paper.⁵

A thesaurus is a dictionary of words and phrases, grouped together according to similarities in their meaning. It contains “terms” and explanations (or “scope notes”) of how those terms should be applied, and it defines relations between the terms. *Which* of these terms, explanations, and relations one selects depends on the intended use(s) and audience(s) of the thesaurus. For detailed retrieval of information from the scientific literature a large and complex thesaurus will be needed,⁶ but for our goal of unlocking web resources for a very diverse audience, a very much simpler approach (targeting the proverbial intelligent 12-year old) may be taken.

1.1. Towards a prototype multilingual thesaurus application

The work of the ArchTerra Consortium is aimed at reaching the following interrelated goals:

- To establish a short list of terms adequately covering the presently available internet resources for European archaeology, with appropriate scope notes (a glossary) and translations;
- To structure these terms in a thesaurus;
- To implement a browseable user interface to the thesaurus, for submission / retrieval;
- To implement a browseable maintenance interface to the thesaurus, for submitting / discussing / activating new terms, definitions, relations, and translations.

Progress toward these goals is discussed in more detail below. Following a discussion of general design considerations in section 2, the remainder of this note is used to describe the user and maintenance interfaces of the prototype thesaurus application.

2. Design considerations

The goals specified above have been used to provide further parameters for the design of the prototype application. Parameters relating to the content of the thesaurus, the types of relations that will be recognised, the implementation of multilinguicity, and the

use of standards are discussed each in turn below. It should be noted in advance that many desiderata have not been implemented in the prototype – rather, it is hoped that they may feed further and wider discussion and development at a later stage.

2.1. Multilinguicity

The requirement of multilinguicity (or, more properly, language-independence) is relevant to the design of the thesaurus application in two ways. How do we build a language independent system that will allow us to add languages as needed? And how do we ensure that submissions and retrievals will be handled correctly irrespective of the language or script being used? The latter question is a technical one, to do with the problematic representation of various European character sets by web browsing software, which we will not explore any further here. The former question leads off to the deeper issue of what constitutes a translation. True one-to-one mapping of terms in two different languages does not occur very often, and a successful thesaurus application must therefore be able to deal with the existence of multiple – multi-cultural – and conflicting meanings of the same term. Among the practical approaches to consider are:

- a) taking an arbitrary decision, e.g. to abide by the definitions given by some authority, and
- b) ensuring professionalism in the translators (avoiding debates about the quality of a translation).

Although the ArchTerra Consortium’s original plan was to implement the prototype in four languages versions (English, Polish, Romanian and Bulgarian), restrictions in the available time, funding, and expertise meant that only the English version was in fact developed. However, as the prototype is language-independent, the addition of languages should be a relatively simple task.

In order to preserve language independent storage of thesaurus information, three linked database tables are used in the ARGE database structure to store information about relations, terms, and relations between terms. Another three linked tables store information about terms assigned to URLs (see figure 1). Terms, Classes, Relations and Languages are all numerically coded, and “lookup” tables are used to translate these codes into words in the user’s chosen language.

2.2. Standards

As a matter of principle, the thesaurus should, where-ever possible, make use of existing international standards. For example, ISO standards or draft standards for the recording of character sets, country names, and language names can be applied without difficulty. But other standards may not be applicable for a variety of reasons; for example, the ISO date standard does not cope well with the needs of archaeological dating, because it uses hyphens as a wildcard character (so –999 expands to the year 0999 or 1999 or 2999) and does not allow the BC/AD system or negative dates. The best system for numeric representation of dates is the astronomical system, where 0 = 1 BC, -1 = 2 BC, etc (Millard, pers. comm.). It goes without saying that, where such standards differ from generally accepted usage, the standard should be used internally while the generally accepted version is presented to the user.

Existing national standards and draft standards for professional terminology, such as the English MIDAS and INSCRIPTION

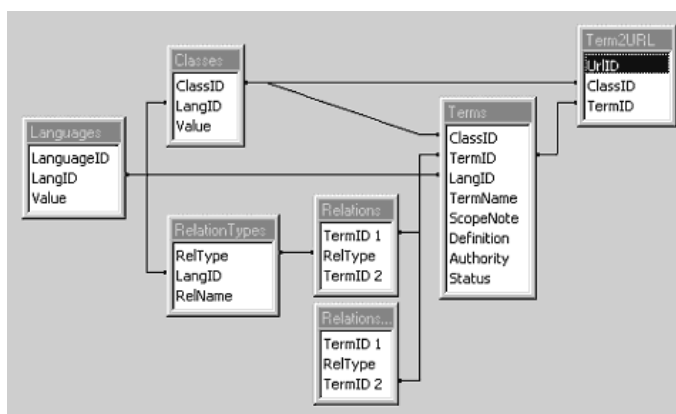


Figure 1: Term specifications are stored in table *Terms*, their relations in table *Relations* (which occurs twice because it creates “internal” links), and their assignment to URLs in table *Term2URL*; Tables *Classes*, *Languages*, and *RelationTypes* are “lookup” tables.

standards, may also be relevant to the thesaurus. MIDAS, the Monument Inventory Data Standard, is a “content” standard for recording architectural or archaeological monuments; INSCRIPTION is a set of standard “word lists” covering things like the Type of a Monument, or terms to describe archaeological periods.

It should be recognised, however, that no effective standards exist in many areas of archaeological terminology, and the creation of such standards is far outside the scope of the ArchTerra Consortium. International professional organisations such as the European Association of Archaeologists (EAA) and the Union Internationale des Sciences Préhistoriques et Protohistoriques (UISPP) should provide the forum for such work.

Worldwide standards also apply to the coding of the interfaces to the thesaurus (see sections 3 to 5), which should operate well across a range of platforms and browsers. In addition to adhering to the HTML standard, the UTF-8 (UNICODE) character set may be used by the interfaces so that western and central European, Baltic, Greek and Cyrillic scripts are all displayed correctly. It should be noted that, while the implementation of multilinguicity *per se* is straightforward, dealing with multiple character sets is not. Although the major browsers now recognise UNICODE encoding of web content, appropriate fonts may not be available to the client software to display all character sets in use across Europe correctly. We can only hope that ongoing standardisation will remove this obstacle before long.

2.3. Content

Despite the relatively restricted goals that we have set for the content of the thesaurus, this is probably the most difficult issue to settle. Relying on the experience gained earlier with ARGE, the ArchTerra Consortium settled on a guideline of including some 100 subject terms, some 30 period terms, and another 100 geographic terms in the thesaurus.

The biggest problem doubtlessly lies in how to avoid/resolve complexities or lack of clarity in the *meaning of* and *relations between* the terms. For example, archaeological periods are defined by a mixture of chronological, cultural, technological, political, and architectural criteria; it is not always clear to which area they

are applicable and what are their absolute start and end dates (if any). Does “eneolithic” indicate the last phase of the Neolithic period, coming after the Late Neolithic phase (as the name implies), or is it a transitional period in its own right (in between the Neolithic and Bronze Age)? The “Roman” period begins and ends at different times in different parts of Europe, and does not exist in others. Can this be dealt with in the Scope note or is a more complex solution called for?

A second problem can be summarised in the question of how to ensure that the thesaurus provides truly equal “coverage” of all subjects, periods, and regions of Europe. For example, terminology for post-Roman periods is generally less precise than that for later prehistory; “non-preferred” terms such as “cave men” may have to be included because these are likely to be used by the public or to occur in older literature; and *local terms* (such as “sub-Appennine” for the Early Iron Age in Italy) may be needed to allow the effective inclusion of many parts of Europe.

Two approaches are available to avoid or alleviate endless disagreements about the content of the thesaurus. Firstly, definitions and scope notes describing the meaning and clarifying the applicability of terms in language appropriate to the audience will, whenever practicable, be taken verbatim from an authoritative source. Scope notes should also specify the amount of “fuzziness” in the meaning of the term. Secondly, the thesaurus will be “coarse-grained” in its description of archaeological internet resources, i.e., many potential conflicts of interpretation will be avoided because specific terms such as “Flavian” will be mapped to more general terms such as “Early Roman Empire”. As a further antidote to the potential confusion, at least in the area of chronological terminology, we can include absolute chronology in our thesaurus by offering *millennia* and *centuries BC/AD* as indexing options (the BP and Cal systems of dating are not familiar to the general user, but could also be added if the need arose).

2.4. Relation types

The following standard thesaurus relations will be recognised initially by the system:

- Class (CL; top-level descriptive category to which the term belongs, e.g. “Chronological Period”)
- Scope Note (SN; a brief statement about the scope of the term which could include a named authority for the definition used, geographic and temporal ranges, etc.)
- Broad Term (BT; indicates the generic term of which the current term is a specific; a.k.a. “is a kind of”)
- Narrow Term (NT; indicates the specific terms of which the current term is the generic; a.k.a. “contains”)
- Related term (RT; indicates all relation types not covered by BT and NT)
- Used for (UF; indicates any alternative but non-preferred terms with the same meaning; a.k.a. “is a preferred term for”)
- Use (USE; indicates which other term of the same meaning is preferred over the current term; a.k.a. “is a non-preferred term for”)

This can be extended at a later stage with relation types that are more specific to archaeology, e.g., Precedes (P; i.e. chronologically) and Succeeds (S; i.e. chronologically).

The importance of defining the relation types themselves may be illustrated by an example. The ISO standard for monolingual thesauri recommends that the use of BT and NT for part/whole relationships should be restricted to a few specific types of term: (a) systems and organs of the body; (b) geographical locations; (c) disciplines or fields of discourse; (d) hierarchical social structures. This would allow a BT/NT relation to exist between (a) skull and lower jaw, (b) Scandinavia and Norway, (c) Geophysics and GPR, and (d) tribe and moiety, but it would not allow (e) arms and sword.

One useful feature is the fact that terms need not be fully “linked up” to other terms in the thesaurus. For many colloquial search terms (that is, terms that non-specialist users would like to use) a simple mapping to one or more of the existing “preferred” terms in the thesaurus should be sufficient for effective retrieval. The thesaurus application would contain only the non-preferred term, its scope note, and its USE relation (to a preferred term), and any searches would take place “behind the scenes” using the preferred term. The same solution could be applied to the use of *local terms* as mentioned above, although here the relation type would be RT (related term) rather than USE.

3. The user interface

End users of the thesaurus will want to do two things with it – retrieval and submission of web resources. Logically, retrieval with the help of a thesaurus is just one among several search mechanisms, and the thesaurus option should therefore be offered in the “Search Page” of a web site. To keep the user’s retrieval interface efficient and effective, the thesaurus must be activated in two stages. Primary retrieval of archaeological records should be achieved through a limited and hierarchical (i.e. using only the BT/NT relations in the thesaurus) interface, implemented either as a series of foldout menus or as a collapsible, Windows Explorer-like menu such as that used by the NAVIS web site. The results of the primary search will be accompanied by dynamically generated options for widening or narrowing searches which constitute the secondary retrieval system.

The code for the primary retrieval menu, presenting the BT/NT part of the thesaurus, may have to be periodically generated offline in order to prevent the application from becoming unacceptably slow and complex (many menus with many options in the user interface). The “root” menu would contain CLASSES such as *period*, *region*, *organisation type* and *data provider*, each with no more than 10 options. The peculiar nature of the user community – a mixture of old and young, lay and professional – puts specific requirements on the design of the user interface. While users should be made aware that their actions are “filtered” by the thesaurus application, the presence of the thesaurus should be non-obtrusive and unnecessary use of jargon should be avoided. Since archaeological terminology is neither consistent nor logical, a heavy burden rests on the *definitions* and/or *scope notes* attached to the terms. These should therefore be accessible to the user at all times. Contextual information about each term can be presented to the user by employing an “onMouseOver” method (JavaScript) to show the scope note for each term; the terms selected by the user should also be echoed to the screen so that the user keeps track of

his/her actions. A thorough study of interface design will be needed to upgrade the current prototype to a working “alpha” version.

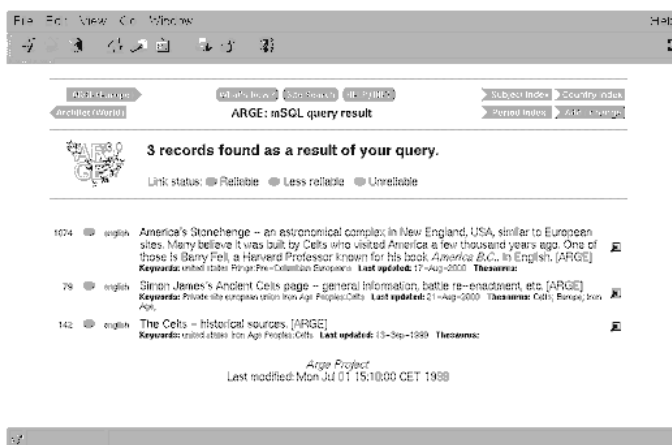
The secondary retrieval menu (part of the output page resulting from a particular thesaurus query) will allow users to refine or widen their queries by hyperlinking each key term of the currently selected web resources to its thesaurus relations (see figure 2). Whilst a direct user search using the thesaurus has not been implemented at this date, a demonstration of the secondary retrieval system was implemented, and is shown below. A preliminary series of hierarchical relations (NT, BT) have been defined between the terms on the shortlist for testing purposes.

All of the above assumes that the user will be happy to use only the terms offered by the thesaurus for searching, but this is an unlikely scenario. Could the thesaurus application be taught to understand free-form keywords submitted by the user? Are there others ways in which the user can be shielded from the very structured search environment offered by a thesaurus? Among the potentially useful tools and techniques that can be explored are “mapping” terms or sets of terms to others, parsing the user’s search strings to allow orthographic and syntactical variation, and searching by example:

- One of the most useful concepts in the enormously complex issue of creating an acceptable thesaurus for European archaeology is that of “mapping”. A multiplicity of period terms deriving from cultural, political, chronological, and architectural criteria, for example, can perhaps be mapped to the single “yardstick” of absolute dates. “Flavian” would map to AD 68 <> AD 95 (or whatever), which would map to “Roman” in France but to “Late Iron Age” in Denmark. The implications of such mapping must be investigated. What happens if such dates change? How should overlapping periods be handled? Mapping can also be used to link two or more thesauri together, so that, for example, users can use the familiar terms from a local archaeological thesaurus for constructing queries which will be “translated” to equivalent terms in the European thesaurus before execution;
- A parser, to resolve orthographic and syntactical variation, including errors, in search strings. Many search engines already implement this as “near” searching. One simple approach would be to search case-insensitive and to ignore everything that isn’t a letter, e.g., /post[^a-z]*media?eval/i is an expression that will match most of the possible spellings of that term, including the American ones;
- Searching “by example”. While a lay public may not be able to handle archaeological terms well (e.g., being unaware that they should ask for “Palaeolithic” to get information about Neanderthals), this would not be a problem if they could “search by example”. The thesaurus could include a list of likely “examples” (the Iceman, Stonehenge) for this purpose, and “map” these to one or more proper archaeological terms.

4. The maintenance interface

The maintenance of our prototype thesaurus application involves two separate issues: the maintenance of the thesaurus itself (adding, editing, or deleting terms and relations), and the indexing of web resources using that thesaurus. The following sections re-



a)



b)

Figure 2: (a) A user query on the string “celts” results in 3 URL records being displayed by ARGE. URL 79 (Simon James’s Ancient Celts page) has attached thesaurus terms “ ”. (b) When the user clicks on any thesaurus term, a helper window displays the current thesaurus relations of that term; when the user selects any of these, a new query is executed and new results are written to the main window. This cycle can be repeated at will.

view both these issues, using examples from ARGE’s prototype thesaurus maintenance interface.

4.1. Indexing web resources

Once the thesaurus itself is complete and available, the main task awaiting the maintainer of a web guide is to index new URLs, and to check/edit indexed URLs submitted by users. The indexing interface will present all currently available information on any URL, with options to add/change/delete key words and to visit the URL itself in a separate window. The regular maintenance interface of ARGE, which already allows key words to be added to a URL record, was extended with options to add/change/remove any number of terms from any number of CLASSES (figure 3). An option to call up the helper window containing the current thesaurus terms and descriptions was also added.

4.2. Building and maintaining the thesaurus

Building and maintaining the thesaurus itself is a task that should be restricted to an editorial board affiliated to an appropriate pro-

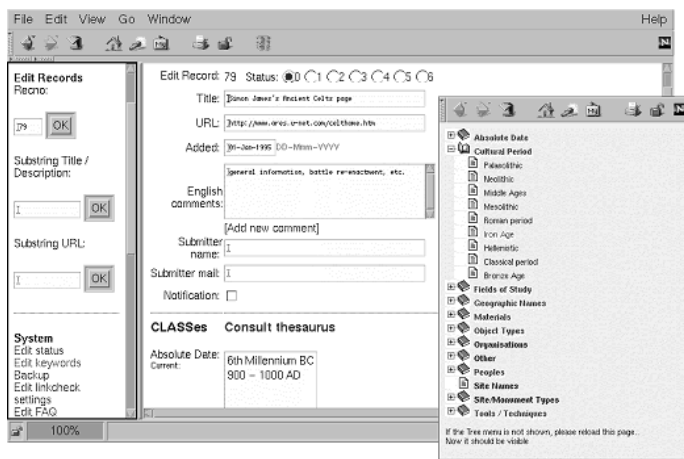


Figure 3: The URL editing page of ARGE's maintenance interface. URL 79 (Simon James's Ancient Celts page) has been selected for editing in the main frame. Any number of terms may now be selected from dynamically generated option lists for each currently defined thesaurus CLASS. Clicking on the update button at the bottom of this form updates the record across all database tables.

professional organisation. The maintenance interface should have the following functionality:

- Authority to perform actions should be stored in user profiles;
- An Editor must be able to view current relations between terms, to add and edit terms, scope notes, relation types, and relations between terms; the use of definitions in the scope note requires the addition of an "authority" field;
- A Translator should be able to add and edit translations.

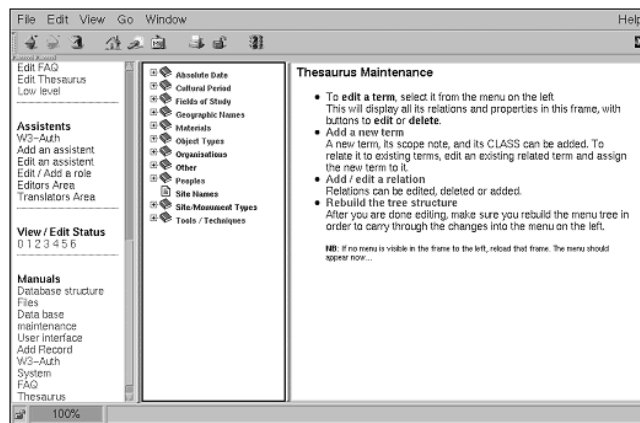
The prototype thesaurus maintenance interface has been implemented as a "page" in the regular maintenance interface of ARGE (see figure 4). While the right hand frame provides links to the forms needed for adding terms to, and removing them from, the thesaurus, and for adding/editing/deleting a thesaurus relation or relation type, the left hand frame presents the hierarchical part of the current thesaurus. When clicked on an icon, the tree expands to its next level; when clicked on a term its relations are displayed in the right hand frame for editing.

5. Conclusions and further work

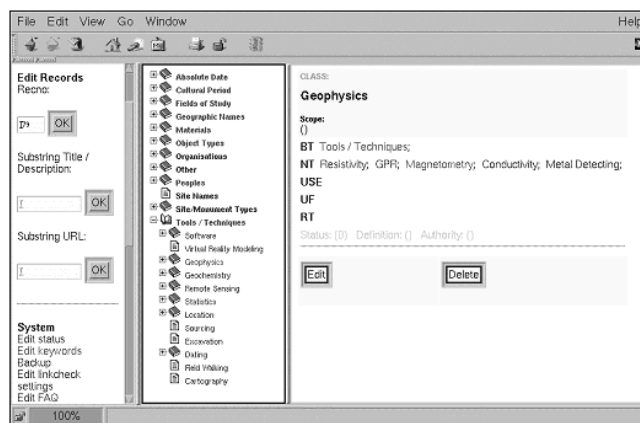
The exponential growth in the amount of archaeological data being published on the Internet has in recent years brought to the fore a more professional and structured approach to retrieving and indexing web resources. The use of a thesaurus to structure a body of information, as proposed and developed in prototype by the ArchTerra Consortium, is just one example of the way in which "metadata" are likely to transform our personal and professional use of the web.

The discussion, design and development of the prototype, of which this paper is intended to be a "snapshot", have brought to light a large number of issues which will need a follow-up. In no particular order these are:

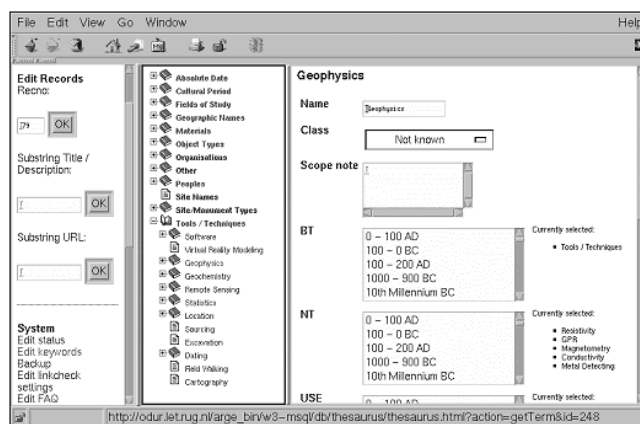
- Speed and Stability.** The prototype was developed on an HP 9000/180 workstation under a CERN HTTP server and



a)



b)



c)

Figure 4: Thesaurus maintenance page of ARGE (a), leading off to forms for adding or editing terms (not depicted) and for viewing (b) and editing (c) or deleting relations between terms.

Hughes mSQL 2.09; a more stable and fast platform will be needed for the development of an alpha version.

- Content.** Although the database structure underlying the thesaurus application is now in place, its current content is temporary and incomplete. In particular, relations between terms have only been inserted for testing purposes so far, and all content is in English. A priority for further work is therefore to complete the short (~250 term) thesaurus envisaged originally by the ArchTerra Consortium, and to verify that the prototype works in at least one other language. The medium of online discussion lists has been found to work well for this sort of work, which should

preferably take place under the auspices of a professional organisation such as the Association of European Archaeologists. If an editorial board is established, it should have a forum and archive for discussing thesaurus maintenance, and for receiving proposals for new terms and relations from users. This implies that some sort of status tracking should also be implemented.

- *Interface design.* Little effort has gone into the on-screen layout of the prototype, and further work will have to include serious investment in studying the requirements of effective interfaces. The thesaurus environment is unlikely to be familiar to the general public, so extra care must be taken to remove unnecessary jargon, complex layout, etc.
- *Natural searching.* Among the potentially useful tools and techniques that can be explored to create a more natural user interface are “mapping” terms or sets of terms to others, parsing the user’s search strings to allow orthographic and syntactical variation, and searching by example.
- *Trust and Authority.* Many archaeologists have expressed concern with the “power” wielded by an editor of a web guide and, by implication, of the editor of a thesaurus. The gist of my reply to such concerns (Van Leusen et al. 1996) has been to note that it is not a matter of power but one of trust or authority; any editor (whether traditional or online) has to earn, and can lose, the user’s trust by the quality of his or her work. ARGE has piloted a system of remote online editors whose contributions are tagged with their identity, so that users can evaluate their authority if they wish to do so. Optionally, the quality of the editors could be further safeguarded by a professional accreditation system, for example with the Association of European Archaeologists (EAA).

Acknowledgements

In writing this paper, and especially the “design considerations” section, I have profited hugely from the discussions in the FISHEN e-conference on building a period thesaurus for the UK (January 2000); I have attempted to acknowledge all significant re-use of the ideas and formulations of the participants in footnotes but here would like to acknowledge their contribution as a group too. Previous versions of this paper were reviewed during another online workshop, this one organised by the ArchTerra Consortium (April 2000), and a workshop during the CAA2000 conference (Ljubljana, April 2000). Among the participants at the latter workshop, I

would like to thank Martin Doerr especially for his useful presentation and comments. Further thanks go to Allard Mees for his regular e-mailed questions and criticism.

References

- CHAMPION, S., 1997. Archaeology on the World-Wide Web: a user’s field-guide, *Antiquity* 71: 1027-38.
- INTERREG II C project “Co-operation in spatial planning” (2000-2001, no web site available yet)
- TERM-IT project “Multilingual support for multimedia services” (1998-1999, <http://www.mda.org.uk/term-it/welcome.htm>)
- VAN LEUSEN, M., CHAMPION, S., LIZEE, J. and PLUNKETT, T., 1996. Toward a European Archaeological Heritage Web. In Kamermand, H. and Fennema, K. (eds.), *Interfacing the Past. Computer Applications and Quantitative Methods in Archaeology CAA95* (Analecta Praehistorica Leidensia 28), pp. 507-15. Leiden: University of Leiden.
- WILL, L., 1999. Software for building and editing thesauri (<http://www.willpower.demon.co.uk/thessoft.htm>)

Notes

- ¹ The World Wide Web Virtual Library for European Archaeology, at <http://www.let.rug.nl/arge/> (since 1995).
- ² For more information about this, go to <http://www.purl.org>.
- ³ 996 - 1998, EC Telematics 2c program, at http://www.prosoma.lu/cgi-bin/show_new.py?id=4482&page=description and <http://www.inria.fr>.
- ⁴ For more information about this, go to <http://archterra.cilea.it>.
- ⁵ For further information about the prototype, including updates, go to <http://www.let.rug.nl/arge/Work/thesaurus.html>. The thesaurus prototype was developed at the University of Groningen by Robert Bouma, based on the mSQL database (version 2.09) and Perl (version 5).
- ⁶ See, for example, the RCHME thesaurus of Monument Types and the MDA thesaurus of archaeological objects (both at http://www.rchme.gov.uk/thesaurus/thes_splash.htm), the Bronze Age glossary produced by the Bronze Age Campaign project (Council of Europe, at <http://www.cimec.ro/arh/bronz/>), and the January/February 2000 archives of the FISHEN e-conference (at <http://www.mailbase.ac.uk/lists/fishen/archive.html>). For the “position paper” by Jeremy Oetgen which started off the FISHEN debate, see Appendix 1.

Appendix 1: Chronological terminology in the UK and Ireland

Appendices were written by Jeremy Oetgen, British and Irish Archaeological Bibliography, c/o The British Academy, 10 Carlton House Terrace, London SW1Y 5AH, United Kingdom, e-mail: j.oetgen@britac.ac.uk

Reviewed with reference to the British and Irish Archaeological Bibliography and the MIDAS standard.

Adapted from a deposition sent to the FISHEN meeting on 22 November 1999.

This contribution illustrates the diversity of chronological frameworks in use in archaeology in the British Isles and emphasises the need for an integrated scheme for the main periods that is explicit about the assumptions and evidence on which the chronology is based. It is suggested that a standard reference chronology should be set up with the participation of the whole archaeological community, along with those in related disciplines.

1. A conundrum

Did you know that both Julius Caesar and William the Conqueror landed in the south coast of England in the same period? Would it surprise you that period was the Iron Age?

2. An explanation

The RCHME periods list defines the start of the Romano-British period as AD43 - the Claudian invasion. Strict application of this “terminus post quem” would therefore classify Caesar’s expeditions in 55BC and 54BC as Iron Age.

In England it is usual to accept that the Middle Ages proper began in 1066, when William defeated King Harold at Hastings. In Ireland, however, it is commonly considered that the Middle Ages did not begin until the late twelfth century AD, when Normans from England established settlements in Ireland, - furthermore, many Irish archaeologists talk of the period up to the Anglo-Norman invasion as Late Iron Age (see for example Laurence Flanagan’s “A dictionary of Irish archaeology”, published in 1992 by Gill & Macmillan). Thus the Norman Conquest of England could be referred to as an Iron Age event.

3. So what?

Traditionally, parallel, relative chronologies have been used in different regions to deal with the archaeological evidence as it exists in different geographical/cultural areas. We have learned to make allowances for this and we all know what we really mean by the terms Neolithic, Bronze Age, Iron Age etc, don’t we?

But this relativism seems increasingly less rational as modern methods now enable archaeological evidence to be dated in much more absolute terms. Not only are scientific methods - such as radiocarbon determination and tree-ring analysis - providing ever more calendar dates with which to peg down chronologies, but increasingly sophisticated studies of artefacts are refining our ability to date occupation from cultural evidence. Integration of regional chronologies provides better opportunities for identifying cultural parallels and facilitates the study of interregional comparisons.

And, from an information specialist’s perspective (as the example above shows) it cannot be assumed that we all know what we mean by the names we ascribe to archaeological periods - computers certainly don’t!

At the British and Irish Archaeological Bibliography (BIAB) we analyse the content of publications relating to all aspects of archaeology from across the UK and Republic of Ireland; these publications emanate from a wide range of institutions of varying status - both professional and amateur - and many originate from abroad. We classify references according to topic, with the period being of primary importance - because experience tells us that is what people most often look for when searching for archaeological information. At present, when we ascribe a period classification to an article we need to take into account the different chronologies in use in the British Isles and consider carefully how researchers are going to expect to be able to retrieve information. If an Irish archaeologist describes a site as Iron Age, are they saying it dates within a range of c250BC to cAD1170? If so, is an English archaeologist searching for Iron Age references going to be interested in this site? Should a site in Orkney dated to AD200 be classed as “Roman”, “Iron Age”, or something else?

What if new evidence emerges to redefine the dating of a period? The discoverer of that evidence may consider it as incontrovertible and amend their own use of the period name immediately. It may be some time before the rest of the archaeological community learns of the new evidence or realises its significance and many will continue to use an old designation. Personal experience provides me with an example of this: at BIAB we had been using a scheme that gave the start of the Palaeolithic as 350,000BC; however, while reviewing the scheme as a prelude to this FISHEN discussion we realised that the recent hominid find from Boxgrove in Sussex had effectively pushed back to 500,000BP the earliest evidence of human occupation (and thus the start of the Palaeolithic) in Britain.

The Boxgrove example also illustrates how modern dating methods have generally transposed the prehistoric periods back in time. Early archaeologists and antiquaries conceived of the prehistoric past as being much more recent than we now know it was - and that means we need also to consider how our intellectual predecessors constructed their chronologies.

4. Proposal for the creation of a standard reference chronology defining archaeological periods

I hope that what I have just said highlights the need for a standard reference resource that provides authoritative definitions of the key periods (respecting regional and international needs). This resource should also be informative, identifying and explaining the evidence on which the dating is based. To be authoritative the resource must be inclusive of all main chronological systems used in the UK and, preferably, the whole of Ireland.

I would like the reference chronology should include dates for:

- Documented events, e.g. battles/invasions; biographical dates; kingships; documented inventions.
- Periods of currency of monument types and artefact types, certainly for those most representative of their respective periods.

- Cultural and technological innovations, e.g. introduction of metallurgy.

5. Who would use the reference chronology?

A standard archaeological reference chronology would be a valuable tool for a wide range of users (professionals, amateurs, and school and university students) and it ought potentially be of interest to researchers in related disciplines and have a variety of applications. The reference chronology should therefore be made available in a form and medium that would facilitate its use by the widest range of users.

6. Practical procedures for developing the reference chronology

I envisage that the reference chronology would be developed in the following five stages:

1. Collect information on the chronological schemes currently in use by archaeologists in the British Isles. (Out-of-date schemes might also be included for the benefit of anyone working on retrospective data.)
2. Collect any additional chronological data that are relevant (kingship lists and dates for key historical events etc)
3. Collate the received schemes

4. Establish and agree the defining characteristics of each period, based on collation of the received schemes.
5. Define the date ranges for each main period by encapsulating the accepted dates for all the component characteristics.

To maximise the cross-disciplinary potential of the reference chronology the participation should be sought of historians, palaeoclimatologists, environmental historians and experts in other related fields. Data on various chronological schemes could be collected and entered onto a simple database, a possible field structure for which is set out below.

7. Suggested field structure for recording chronological data

Fields are set out below. TPQ = terminus post quem [point after which - i.e. earliest possible date]; TAQ = terminus ante quem [point before which - i.e. latest possible date]. The dates are given in ISO format. Day and month should be included to allow for the possibility of these being known (however remote this may be!) Zeros could signify blank in these fields - although too many noughts could encourage typos! The note fields allow for source to be cited, but it might be worth setting up a separate table of sources. Both TPQ and TAQ are required for Start and End dates

Term	Code Date	Max Date	Min Order	Chronol. Uid	Term Uid	Parent Term
LOWER PALEOLITHIC	LPA	0	0	1	54	5
EARLY PREHISTORIC	EPR	0	0	1	95	0
LATER PREHISTORIC	LPR	0	0	1	94	0
PALAEOLITHIC	PA	0	0	1	5	0
PREHISTORIC OR ROMAN	UP	0	0	1	26	0
PREHISTORIC	PR	0	0	1	52	0
MIDDLE PALAEOLITHIC	MPA	0	0	2	55	5
UPPER PALAEOLITHIC	UPA	0	0	3	56	5
EARLY MESOLITHIC	EME	0	0	4	57	6
MESOLITHIC	ME	0	0	4	6	0
LATE MESOLITHIC	LME	-4000	-7000	5	58	6
NEOLITHIC	NE	-2200	-4000	6	7	
EARLY NEOLITHIC	ENE	-3000	-4000	6	59	7
MIDDLE NEOLITHIC	MNE	-2700	-3500	7	60	7
LATE NEOLITHIC	LNE	-2200	-3000	8	61	7
EARLY BRONZE AGE	EBA	-1500	-2500	9	62	8
BRONZE AGE	BA	-700	-2500	9	8	
MIDDLE BRONZE AGE	MBA	-1000	-1600	10	63	8
LATE BRONZE AGE	LBA	-700	-1000	11	64	8
IRON AGE	IA	43	-800	12	9	0
EARLY IRON AGE	EIA	-400	-800	12	65	9
MIDDLE IRON AGE	MIA	-100	-400	13	66	9
LATE IRON AGE	LIA	-43	-100	14	67	9
ROMAN	RO	410	43	15	10	0
EARLY MEDIEVAL	EM	1066	410	16	11	0
EARLY MED. OR LATER	UM	0	0	16	53	0
MEDIEVAL	MD	1540	1066	17	28	0
POST MEDIEVAL	PM	1901	1540	18	16	0
MODERN	MO	0	1901	19	24	0
UNCERTAIN	UN	0	0	20	25	0

respectively in case of, say, the end of the Roman period where there is debate about how long the Roman way of life persisted.

Record fields

Name (of period/historical event):

Note on name:

Start TPQ: (ISO standard format YYYYMMDD)

Note on Start TPQ:

Start TAQ: (ISO standard format YYYYMMDD)

Note on Start TAQ:

End TPQ: (ISO standard format YYYYMMDD)

Note on End TPQ:

End TAQ: (ISO standard format YYYYMMDD)

Note on End TAQ:

Illustrative examples of completed records

Name: Bronze Age

Note on name:

Start TPQ: -23000000

Note on Start TPQ: established proprietary convention

Start TAQ:

Note on Start TAQ: none

End TPQ:

Note on End TPQ: none End TAQ: -7000000

Note on End TAQ: established proprietary convention

Name: Romano-British

Note on name: period of Roman occupation of Britain

Start TPQ: 00430000

Note on Start TPQ: literary source refs...

Start TAQ:

Note on Start TAQ:

End TPQ: 04100000

Note on End TPQ: literary source refs...

End TAQ: 04500000

Note on End TAQ: assumes persistence of local Roman administration

Appendix 2: Building a European thesaurus for archaeological periods

Many organisations have their own practice and “standard” when it comes to archaeological dating. The Royal Commission on the Historical Monuments of England provides the following list (in use in the English Heritage National Monuments Record database), which dates to 1993 and includes absolute dates and BT/NT relations (© RCHME 1993).

These and other existing documents were discussed in the FISHEN e-conference, following which a draft thesaurus of period terms for the UK was created, of which the following is a small extract (© English Heritage 2000). After review this thesaurus will become part of the UK “Inscription” standard.

Deverel Rimbury Culture	USE MIDDLE BRONZE AGE	LATE MESOLITHIC
EARLY BRONZE AGE	UF Beaker Culture	SN The later phase of the Mesolithic, characterized by the use of narrow blade industries, with smaller and more varied microlith forms. It was during this phase that Britain was physically separated from continental Western Europe.
Wessex Culture	SN The first phase of the Bronze Age, associated with the more widespread use of round barrows, and the first manufacture and use of copper and bronze tools, weapons and ornaments.	CL CULTURAL PERIOD
CL CULTURAL PERIOD	BT BRONZE AGE	BT MESOLITHIC
EARLY IRON AGE	UF Hallstatt Culture	LATE NEOLITHIC
CL CULTURAL PERIOD	BT IRON AGE	UF Beaker Culture
EARLY MEDIEVAL	CL CULTURAL PERIOD	SN The final phase of the Neolithic, associated with the construction and use of henges, and increasingly diverse (and regional) styles of pottery, lithics, etc.
NT ANGLIAN	ANGLO SAXON	CL CULTURAL PERIOD
NORSE	SAXON	BT PALAEO-LITHIC
RT 10TH CENTURY AD	5TH CENTURY AD	LOWER PALAEO-LITHIC
11TH CENTURY AD	6TH CENTURY AD	UF Acheulian
12TH CENTURY AD	7TH CENTURY AD	SN The earliest phase of the Palaeolithic, characterized by the use of bifacial tools, such as handaxes.
13TH CENTURY AD	8TH CENTURY AD	CL CULTURAL PERIOD
14TH CENTURY AD	9TH CENTURY AD	BT PALAEO-LITHIC
15TH CENTURY AD	CL CULTURAL PERIOD	Maglemosian
EARLY MESOLITHIC	SN The earlier, post-glacial phase of the Mesolithic, characterized by the use of broad blade industries, and showing clear continuity from the preceding final stages of the Upper Palaeolithic.	USE MESOLITHIC
		MEDIEVAL
		UF Middle Ages
		CL CULTURAL PERIOD
		RT 11TH CENTURY AD
		12TH CENTURY AD
		13TH CENTURY AD
		14TH CENTURY AD
		15TH CENTURY AD

Even if all countries of Europe were to agree on such nation-wide standards (and not just for period terminology but also for other classes of archaeological terms), the next step of merging such professional thesauri across national, cultural, and language borders will take a huge effort by specialists. The ArchTerra Consortium has undertaken to explore these issues using the following restricted set of period terms for eastern Europe.

ID	Term
1	Ariuşd
2	Aurignacian
3	Basarabi
4	Biskupin
5	Bodrogkeresztur
6	Boian
7	Bronze Age
8	Cucuteni
9	Culture
10	Dacian-Getian
11	Gumelnița
12	Hallstatt
13	Hamangia
14	Iron Age
15	Karanovo
16	La Tene
17	Mesolithic
18	Monteoru
19	Neolithic
20	Otomani
21	Palaeolithic
22	Tiszapolgar
23	Urnfield Culture
24	Vinca
25	Vucedol