Comunicazione Multimediale per i Beni Culturali

Multimedia Communication for Cultural Heritage

a cura di – edited by

FRANCO NICCOLUCCI
SORIN HERMON

Indice – Table of contents

The ArchTerra Project: Extending the European Archaeology Web over Bulgaria, Romania, and Poland

Martijn van Leusen*, Andrzej Prinke**

*Groningen Institute of Archaeology, The Netherlands
**Archaeological Museum, Poznan, Poland

Abstract. *On the 29th of September 1997, the ArchTerra Consortium submitted proposal 977054, entitled "ArchTerra: Extending the European Archaeology Web over Bulgaria, Romania, and Poland" to the European Commission under the INCO Copernicus programme. The proposal was initially placed on a reserve list but was subsequently accepted for funding by the European Commission on the 13th of November 1998. The ArchTerra project ran for a period of two years beginning in January 1999 (fig. 1). This paper describes the project and its most significant acheievements.*

Sommario. *Il 29 settembre del 1997 il Consorzio ArchTerra invió alla Commissione Europea la proposta 997054 intitolata "ArchTerra: Estendere lo European Archaeological web alla Bulgaria, Romania e Polonia", nell'ambito del programma INCO Copernicus. La proposta fu inizialmente posta in lista d'attesa ma successivamente il finanziamento fu accettato il 13 novembre 1998. Il progetto ArchTerra è così stato attivo per un periodo di due anni con inizio nel gennaio 1999 (fig. 1). Quest'articolo descrive il progetto e i suoi risultati più significativi.*

Figure 1. Logo of ArchTerra project

1 Introduction

1.1 Aims and objectives

The networking and internationalisation of existing academic resources, historical and cultural achievements, as well as the research efforts of the academic community in the field of Archaeology using computer, information and communication technologies is a recent trend with a rapidly growing impact on archaeological research, management and education. Since the mid-1990s, Internet services have completely changed the way scientists work: e-mail, usenet, FTP and finally WWW architecture now allow scientists to collaborate on distributed resources, e.g. databases and exhibitions over the WWW. Although countries of Eastern Europe have an important place in the historical and cultural development of Europe, they could not participate substantially in this process of intensifying information exchange, due to their past closedness and current economic problems. There is thus the spectre of an ever widening 'information gap' between the affluent countries of western Europe, and those countries struggling at the margins.

The ArchTerra Consortium has aimed to help redress the imbalances in access to European networking facilities for professional archaeologists from Bulgaria, Romania and Poland, and to provide the impetus for an active expansion of archaeological Internet communication and information services both within CEC (Central European Countries) and between EU and CEC. Accordingly, the objectives of the Consortium were:
- To establish the technical infrastructure and software tools needed to allow researchers in the field of IT in Archaeology from CEC to join the EAW (European Archaeological Web), in the form of national WWW hosts of the ArchWeb network in the three participating CEC (Bulgaria, Romania, Poland). These hosts were to be located at the main research organizations responsible for archiving, maintenance and supply of information in these countries.
- To provide practical demonstration of the trans-national nature and urgency of archaeological research and management, and the benefits and efficiencies of Internet use, to professional and general users alike. End users were to be able to access both the presently available on-line electronic resources and a core set of demonstration resources from CEC (including Web pages, museum databases, live presentations and virtual tours).
- To strengthen existing scientific relations between EU and CEC and to foster long-term joint initiatives for collaboration, demonstrating the richness and fragility of the European archaeological heritage, by bringing together partners and collections from across Europe. To this end, solutions to specific hurdles to international collaboration (translation schemes for languages with different alphabets, multilanguage and multicultural thesauri of terms and articles, international heritage legislation) were to be explored.

A subsidiary objective of the Consortium has been to publish its activities and results as widely as possible, in order to enable all segments of the European archaeological community to sample the deliverables on offer. To this end, publication was to be web-based, by multimedia CD-ROM, and by scientific publication.

1.2 Background

The ArchTerra project developed out of a number of earlier projects in which the Consortium co-ordinators participated, and it was influenced by EC projects which ran previously or concurrently. This background will be briefly sketched here.

Attempts to harness the Internet for communication and information about European archaeology dates back to the early 1990's, when one of us (Van Leusen, then at the University of Amsterdam) began a guide to archaeological Internet resources using *the gopher protocol*. In 1994 Van Leusen teamed up with Sara Champion at Southampton University to produce the first web version of this guide, which became known as ARGE.

Soon it became clear that the guide's content could only be managed efficiently if it were served from a database rather than as a series of static web pages, and that ARGE's functions were being duplicated by very similar guides maintained elsewhere. It was decided to team up with the maintainers of two of these guides in a UKOLUG – funded research project named BAUD (for *British Archaeological URL Database*) which sought to address these two problems. At about the same time ARGE also became a partner in the ArcheoNet European research network, from which its maintenance and several improvement to its interfaces could be part funded.

The participants in the ArchTerra Consortium were computer scientists and archaeologists from four countries of the European Union (The Netherlands, United Kingdom, Italy, and Germany) and three countries of Central and Eastern Europe (Bulgaria, Romania and Poland).

ArchTerra is not a project for conducting new research, it is rather aimed at implementation of already agreed standards and concepts. Here especially important is the work of the International Committee for Documentation of the International Council of Museums (CIDOC – ICOM), which resulted in *International Guidelines for Museum Object Information: The CIDOC Information Categories*, as well as some of the previous EU funded projects in the cultural domain, such as Acquarelle.

2 The Work

2.1 Project management

The general methodology adopted by the ArchTerra project is that of a *non-interleaving workflow of separate work packages,* implemented by research teams established by the project partners especially for this project, a methodology also known as "Communicating Sequential Process" or CSP. T
he project requires six different groups of tasks to be completed:
- project management (meetings, reviews, reports);
- installation of infrastructure (hardware, software, networks);
- web host implementation (server installation, content creation, service implementation);
- museum information system implementation (design of templates, prototype implementation, testing);
- research (workshops, visits); and
- dissemination and exploitation (international conference, CD-ROMs, guides and manuals).

2.2 Infrastructural facilities

The objectives of the Consortium under Work Package 2 were:
- to establish and/or extend existing LANs at the CEC partners' offices with the aim of ensuring Intranet access to each CEC partners' digital museum holdings;
- to establish and/or extend existing external (dialup) facilities with the aim of creating Internet restricted access to each CEC partners' digital holdings and limited hosting services within each national ArchWEB;
- to provide portable systems for the creation, demonstration, and dissemination of the project's digital deliverables;
- to provide stability to the project's Internet presence.

2.2.1 Hardware, software & network infrastructure
The general aim was to establish and maintain both Internet services and database services. The main concerns of the Consortium have been to make as much use as possible of the existing software and expertise at the partners' organisations, to ensure a basic level of compatibility between the partners' systems. A subsidiary concern was to keep the costs of the deliverable scripts and applications down to encourage their eventual uptake by third parties.

2.2.2 Standards for Information Interchange
All project deliverables intended for use on the Internet or an IntraNet conform to the following basic Consortium guidelines regarding information exchange, digital file formats and rights management. The exchange of information between DBMS adheres to the RDF (Dublin Core) metadata scheme and uses XML for its syntax.
The following digital file format standards apply:
- UNICODE for text encoding;
- JPEG for still images; originals are to be stored at a minimum resolution of 300 dpi and 24 bit color depth; lower resolution and reduced color versions are allowed for use within the MIS;
- MPEG for video files; MP3 for sound files;
- QuickTime VR for 3-D representations; VRML for 3-D reconstructions.
Control over rights in database applications is managed by a system of user profiles. The database systems managed by the Consortium partners have modules for the definition of such profiles.
The following general set of user profiles was defined in accordance with the recommendations made in the mid-term project review:
- *Database Administrator / Superuser* (full rights);
- *Editor* (rights to add/change/delete specified records and fields; eg, museum staff members in the case of the MIS, or subject editor in the case of ARGE);
- *Translator* (special case of Editor, with rights to change only specific database records and fields in a specific language);

- *Client* (special case of Editor of the MIS; with rights only to those records entered by him/herself);
- *Registered user* (privileged access rights; eg, a field project manager in MEDIOLANVM);
- *Anonymous user* (public access rights).

No standards for digitised content other than those proposed in the CIDOC Model for museum information systems (MIS) were adopted by the Consortium.

2.2.3 Evaluation

The installation of hardware, software, and networking infrastructure for the ArchTerra project has generally gone well, despite delays due to problems in submitting cost claims and receiving reimbursements at an early stage. The additional PC's and servers installed by partners MAP, CIMEC, and AIM have extended their existing LANs with the capacity for Internet and Intranet serving of data (in the case of AIM such a LAN was established for the first time), and these systems have been in successful operation at least since the middle of 2000. The laptop computers have been intensively used by the Consortium partners as portable data entry and presentation machines; in combination with the scanning and digital camera equipment they were used to digitise materials for the Museum Information Systems (see below) and the project's web resources (see below), and they were used as presentation machines during Consortium meetings, conferences, workshops, etc.

2.3 ArchTerra's Internet Presence

The Consortium's decision to take advantage of access via Internet was motivated by the great expansion that the Web has had over recent years, the ever-increasing number of its users, the simplicity and practicality of access to resources in Internet, and the great potential that on a daily basis is being developed and transformed into ever more interesting and wide-ranging services, whose usefulness meets the needs of its continually expanding subscribers.

One development in particular has influenced the approach taken by the ArchTerra Consortium: the move from static to dynamic information provision. All the principal manufacturers of data management software have in recent years dedicated considerable resources and attention to the Internet universe, equipping their top-range products in the DBMS (database management systems) sector with modules that permit management of information (databases) within the Internet environment. This commitment has immensely increased the potential of the Web: Internet applications have moved rapidly from a static to a dynamic modality. Alongside the classic Web pages that are conceived, designed and created in static terms, many (perhaps most) pages now provide information content that is dynamically created on the request of the user. The information feeding this type of page is held in DBMS structures and between the two (page and database) a fruitful "collaboration" is set up.

2.3.1 Central Services

The ArchTerra Consortium has undertaken to develop and maintain three 'central' services for the European Archaeological community. The first of these is ARGE, a so-called 'virtual library' storing information about Internet resources for European Archaeology; a prototype web management system was developed in order to facilitate the maintenance of such virtual libraries (figs. 2-4).

The current version, ARGE VLMS 3.0, is available at www.let.rug.nl/arge.
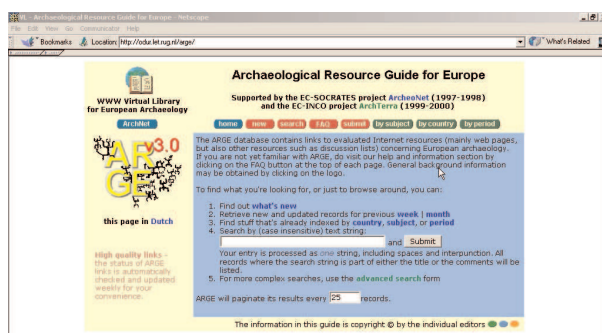
Figure 2. ARGE front web page

Figure 3. Search page of ARGE, by countries
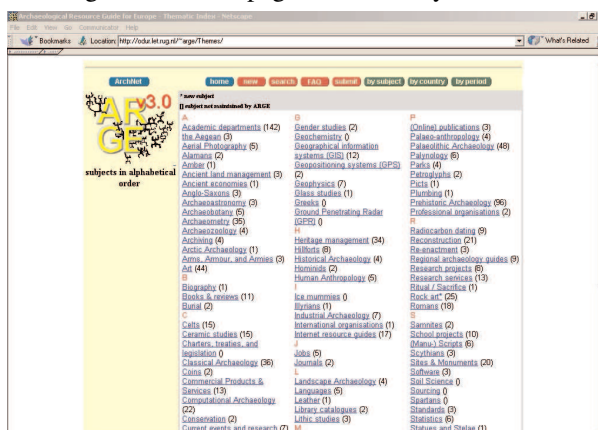
Figure 4. Search page of ARGE, by subjects

Figure 5. Web-page of Mediolanvm service

The second central service for European Archaeology is an internet 'brokering' service for archaeological field work, based on a prototype developed in 1997-8 by the EC-Thematic Network ArcheoNet. This service, named MEDIOLANVM, is available at archeonet.cilea.it/archeosite (fig. 5).

The third central resource developed under ArchTerra consists of a series of glossaries and thesauri for the indexing and retrieval of archaeological objects and information.

ARGE VLMS version 3.0

Under the ArchTerra contract, partner RUG developed a prototype online management system for virtual libraries (VLMS) based on the existing management system of the virtual library for European Archaeology ARGE (Archaeological Resource Guide for Europe). This system was developed for use on a generic UNIX OS using Apache as the internet server and the freeware mSQL as the database server.

The VLMS design and prototype implementation were evaluated by external experts on two occasions; first in September 1999 by Dr Allard Mees of the Römisch-Germanisches Zentral-

Museum in Mainz (Germany); secondly in December 2000 by Dr Paul Tyers, an independent consultant working in the UK. The prototype thesaurus module was evaluated by Dr Leonard Will (UK).

During the ArchTerra project, between October 1998 and February 2001, the daily number of users (hosts) of ARGE grew 6-fold; at the same time, navigation efficiency was doubled so that the average daily number of page requests, 'hits', and MB downloaded has increased only by less than 3-fold. In the same period, almost 400 new evaluated URLs were added to ARGE (a 33% growth), 340 other URLs were updated, and 50 'dead' URLs were deleted. ARGE VLMS 3.0 received the following unsolicited web awards and listings (see www.let.rug.nl/arge/General/prizes.html for details): *Meilleur site 2000* (bonweb.com), listing in MicrosoftPress 2000 Guide of Best Websites, Britannica Internet Guide Award, listing in ISI *Current Web Contents,* Links2Go *Key Resource*, Higly Rated by *Schoolzone*, listing in the UNESCO *Millennium Guide to Cultural Resources on the Web* (CD-ROM with UNESCO World Culture Report 2000).

Mediolanvm

The Mediolanvm web service was created for ArchTerra by the Department of Antiquity Sciences - Archaeological Section of the University of Milan and the Lombard Academic Computing Centre CILEA. Its goal is to facilitate international cooperation in the planning and execution of archaeological field work throughout Europe. It does this mainly by maintaining an easy-to-use web site for information exchange between those who offer and search for field work opportunities. An online database allows students and researchers to find out about, and take part in, archaeological fieldwork projects conducted all over Europe; conversely, it allows European universities and other organisations planning archaeological fieldwork to post full information on the Internet.

The name MEDIOLANVM expresses both its location in Milan and its mediating role in promoting archaeological fieldwork in Europe. The application is driven by an INFORMIX database which serves and personalizes the forms and navigation pages as the user travels through the web site. It has been designed to require minimum maintenance and to automate, as far as possible, communication with the user. A series of forms leads the user through the main steps involved in searching for, or submitting, information on archaeological field work.

The MEDIOLANVM service aims to achieve three main results:
- To facilitate the participation of students and researchers in archaeological research projects in various European countries. The project originated from the need to put people interested directly in contact with European universities or other institutions who are conducting archaeological excavations or surveys.
- To enable those in charge of an archeaological excavation or survey to announce their project and insert it directly onto the Website, so as to offer students and researchers opportunities for jobs or to gain cultural and scientific experience by taking part in archeaological excavations or surveys.
- To place at the disposal of academics and site directors a special type of electronic bulletin board that facilitates communications among these persons and on which to post proposals, researches, suggestions and exchange cultural and scientific information.

MEDIOLANVM offers the following choices:
- To search for archaeological field work using either a geographical search based on clickable maps, or a direct search on chronological period, project type, site type, field work director, teaching level, or project date;
- To apply to take part in a field work project by filling out a web form and sending to the project's director;
- To announce a new archaeological field work project and add it to the database, by filling out a series of forms;
- To exchange news and information regarding any aspect of European archaeological field work on an electronic bulletin board.

Glossaries and Thesauri

One of the core objectives of the ArchTerra project has been to break down existing language barriers to the communication of archaeological information and resources. To this end a Multilingual Interfaces Workshop was held at the joint Computer Applications in Archaeology / Union Internationale des Sciences Prehistoriques et Protohistoriques Commission IV (CAA/UISPP) conference in April 2000 in Ljubljana, Slovenia. It resulted in the *Internet report on the use of*

multilingual thesauri and the archaeological glossaries and thesauri themselves, the latest version of which is available as a Word file from www.let.rug.nl/arge/Work/thesaurus.html. The document, which discusses the reasons for wanting to have a multilingual thesaurus application for indexing and retrieval of archaeological Internet resources, the design guidelines, and the prototype application built at Groningen University, has been reviewed twice. In the early months of 2000 it was reviewed by the members of FISHEN (a UK-centred discussion list dedicated to the creation of a period thesaurus for British archaeology); in April 2000 it was presented and discussed during the workshop mentioned above. It will be published as a paper in the forthcoming proceedings of that meeting.

Deliverable 13 consists of three linked documents:

1. the thesaurus of indexing terms implemented within the Museum Information System of the Archaeological Institute with Museum of Sofia (AIM; English version; the implementation itself is bilingual Bulgarian-English);
2. a multilingual dictionary of archaeological terms, covering periods and cultures in the archaeology of Central Europe. The five languages used by the dictionary are: Bulgarian, Polish, Romanian, Italian, and English;
3. the thesaurus of indexing terms implemented within ARGE Virtual Library Management System 3.0 (English version; the implementation itself is language independent).

Each of these explores different aspects of the international and multilingual indexing and retrieval of archaeological information. Document A, produced by staff at AIM, exemplifies the specialist use of exhaustive restricted hierarchical terms lists (in this case for the classification of the holdings of AIM according to the classes *materials*, *chronology*, and *subjects*). Document B, the result of a collaboration between several Consortium partners co-ordinated by CIMEC, concentrates on problems in the translation and multinational / multicultural use of archaeological thesauri in a first attempt to define a set of terms valid across Central Europe and find equivalents in several important languages. Document C, produced by RUG, represents a more generalist, 'opportunistic', approach to the construction of archaeological thesauri by attempting to organise indexing terms of the Internet resources listed in ARGE, and explores in particular problems relating to the process of construction itself.

Together, these documents provide the materials needed to present the potential benefits and problems of European archaeological thesauri to a wider professional public, to open up discussion, and to prepare proposals for follow-up projects.

2.3.2 ArchWebs

ArchWebs are national nodes for accessing archaeological web resources. Though no formal organisation exists, ArchWebs have been formed and maintained in a number of western European countries, and the ArchTerra project has set out to implement similar nodes for Romanian, Polish, and Bulgarian archaeology. To this end, a document giving guidelines for the structure of an ArchWeb site was produced (deliverable 6), and some software tools were developed in order to be able to provide the functionality required by these guidelines.

ArchWeb tools and guidelines

Since the ArchWeb concept is based on voluntary adherence to a joint navigation structure, the Consortium has decided to produce guidelines, rather than a set of mandatory rules, for the content and structuring of national archaeological gateways of the ArchWeb type which could be used to help the CEC consortium partners build their own ArchWeb nodes. These guidelines were based on an examination of four existing national archaeological web pages – the page of Romanian Archaeology (part of the CIMEC web site), the page of French Archaeology (part of the Archdata web site), the web site of German Archaeology, and the CBA Guide to UK Archaeology online, plus an analysis of the materials proposed by Bulgarian archaeologists to be included in the Bulgarian national ArchWeb. They are being made freely available through the Consortium's web pages to archaeologists elsewhere in Europe who intend to build similar nodes in their own countries.

The full guidelines may be consulted in deliverable 6; here, we will review only its general considerations and design principles.

General Considerations

The structure of an ArchWeb must follow from its objectives. There is no consensus among existing ArchWebs as to what these objectives should be, but all current implementations live somewhere on

the continuum between two extremes – that of the 'gateway' and that of the 'service provider'. An intermediate objective is that of providing an introductory overview of national archaeology. There is usually a trade-off between the level of service you want to provide through an ArchWeb, and the level of maintenance you can afford. This problem may be alleviated by structuring as many services as possible in a self-organising and self-maintaining manner.

Design principles

The following design principles guide the organisation of the national Archwebs in the framework of the ArchTerra project.

A general distinction should be made between the *static* or background information (relating to the existing administrative, research and educational infrastructure, as well as to the existing national and international legislation), and the *dynamic* information concerning current and upcoming projects, events, exhibitions, etc. Such a distinction was, for example, only partially realised in the web page of Romanian Archaeology, due to the subordinate position of the archaeological page to the CIMEC Home page. We consider that the distinction between static and dynamic information should be expressed at the graphical level, too.

A second distinction is to be made between the different *functions* of the ArchWeb: representative (providing official information), informational (providing the possibility for intelligent information retrieval), hosting (of individual and institutional archaeological web pages), communication (a place for information exchange, discussion groups, feedback information), income generation (CD-ROMs, books, space- and subject- limited possibilities for advertising and online shopping). As far as the latter is allowed by the hosting organisation, income may be generated in order to ensure the continuation of the ArchWeb node.

A third, and relatively straightforward, distinction is to be made between meta-information concerning the web site itself – the ArchTerra project, respectively the INCO programme, the partners, etc., and the archaeological information and services it was created to provide. These two different types of information should also be differentiated at the graphical level.

Fourthly, access to dynamic information should be controlled by a system of user profiles, as listed below:

    Database Administrator / Superuser (full rights)
    Editor (rights to add/change/delete specified records and fields)
    Translator (special case of editor, with rights only in a specified language)
    Client (of the MIS; has editors rights only to those records entered by him/herself)
    Registered user (privileged access rights)
    Anonymous user (public access rights)

Finally, as each ArchWeb site will be accessible in at least two languages – English and the national language of the site – both versions should be adapted to the needs of different users types. We consider that an English version has two main objectives: to present national archaeology to outsiders and to allow a certain level of searching and participation by foreign archaeologists (for example in local excavations). This means that the English version must contain most items from the background information, certain dynamic information which is of international interest, and (where applicable) a section concerning special regulations for foreign participation in fieldwork and other research, and for studying archaeology in the country of interest. Where sections of the site are translated, it is very important that professional translations are provided. Specific matters, like legislative texts, could be presented in bilingual format, accompanied by comments both in English and in the local language.

ArchWeb tools

The HTML development tool of choice for the three ArchWebs was DreamWeaver 3. JavaScript applications were used for enhancing the interfaces with pull-down menus, hover buttons, and the like. PHP, ASP and CGI scripts were used for the creation of dynamic pages providing access to underlying MySQL, MS-SQLServer and MS-Access, and FoxPro databases containing information about persons, organisations, events, and archaeological sites and objects.

These tools can be adapted to the needs of further ArchWebs and other archaeological web sites which are developing throughout Eastern Europe, and will be made available by the Consortium partners on request.

ArchWeb-Bulgaria

ArchWeb-BG follows the structural guidelines by the Consortium. It presents contemporary

archaeological investigations in Bulgaria in a thematic manner and gives a concise overview of the major interdisciplinary domains and achievements. Its structure is subordinated to the archaeological periodization schemes used in Bulgarian archaeology. Our aim is to present, in a summary form, the basic development stages of Bulgarian archaeology as a profession. The panel of 15 academic contributors consists of the best known specialists in the respective domains and also includes foreign scientists that have worked or are currently working in Bulgaria. All approximately 100 pages of texts are richly illustrated and are hyperlinked throughout in order to allow fast 'vertical' and 'horizontal' switches within the frame of the presentation of Bulgarian archaeology.

The institutional aspect of Bulgarian archaeology is also presented. This section includes a brief history of archaeology in Bulgaria, a listing of archaeological museums, relevant legislation, current projects, and an overview of the organization of education in archaeology. A special accent is put on the changes that have occurred in the structure of education in archaeology following the political changes of 1989 – e.g. the increasing number of academic archaeology departments in the country. The leading archaeological periodicals and their addresses are also presented.

ArchWeb-BG also provides on-line means, on the principle of equal opportunity, to all professional archaeologists in Bulgaria to upload personal webpages presenting their individual interests and work. The same offer of hosting web pages has been extended to those archaeological organisations which still do not have, or cannot afford, a well-developed Internet presence.

ArchWeb-Poland

ArchWeb-PL was developed according to the guidelines for national archaeology gateways set out by the Consortium (fig. 6). It contains basic data on Polish archaeology, addressing both local and foreign professionals and advanced amateurs; among the resources present are databases on Polish archaeologists and archaeological institutions which contain basic data on all graduated archaeologists that are professionally active (ca 850 records), as well as all institutions which are fully or partly active in archaeological research, education, museology, conservation, etc (ca 230 records). This is followed by a choice of legal acts (full text) that are crucial for archaeological activities as well as for the protection of archaeological sites and monuments.

Another section presents Polish archaeological journals, and contains data on their editors, reproductions of the cover and contents of the latest volume (ca 60 titles, from which 20 are fully presented, to be continued beyond the ArchTerra project). A set of on-line publications is also included in this section, with a choice of the newest, rare and/or important texts, currently including articles on air photos in archaeology, ethics of the archaeological profession, and a richly illustrated album on the results of giant rescue excavations along the Polish part of a trans-European gas pipeline.

Figure 6. Web-page of ArchWeb Poland

Other sections cover conferences and exhibitions – both coming in the near future and already executed, so that it is both a sort of calender for professionals and an archive of past events; a section on Field Research presents excavations carried out by Polish archaeologists at home and abroad, mainly in Africa (Egypt, Sudan); and one page offers links to other related web pages (ARGE, Mediolanvm, EAA) and to the individual home pages of Polish archaeologists.

Because of its significance as a national service for Polish archaeology, a few words must be said here about the dynamic part of ArchWeb-PL. The on-line database of Polish archaeologists and archaeological institutions uses a specially designed textual data format with inserted HTML tags to facilitate data retrieval. Queries are entered from a web page using a form with a JavaScript driven system of menus where the user can restrict the query by institution type and/or administrative district. The request is sent to a script written in Perl which searches the data according to criteria defined by the user and returns a Web document with a list of entries found. In the case of individuals, the data displayed include his/her e-mail address and the name of the institution where he/she is employed; both are hyperlinked so that the user can send a message directly to the person in question or see the details of the institution. The latter data is generated by another script which returns the name, postal address, e-mail and web page (if any) for that institution. Clicking on the institution name link sends another request to still another script which returns a list of all the people working for that institution. If the initial search is performed on institutions, then the first web page generated displays a list of all institution names containing the keyword or string entered in the query form (as restricted by type or area). Alternatively, full lists of people and institutions for a given area can also be accessed from a clickable map interface of Poland, linked to the same set of scripts. The maintenance of the underlying data is shared by the database administrator and the users. The latter may use a web form to add a new individual and/or institution to the database but the process is moderated by the administrator who validates the incoming data.

ArchWeb-Romania

ArchWeb-RO was based on the pre-existing website of CIMEC, the Center for the Cultural Heritage of the Romanian Ministry of Culture (Fig. 7). As suggested by the Consortium guidelines, it has sections providing access to Romanian instititions, publications, and legislation, indexes of journals, and presentations of important archaeological sites: Targul de Floci, Histria, Targsor. The CIMEC databases of archaeological sites and of archaeological excavations in Romania 1983 – 2000 can also be accessed from ArchWeb-RO. The current number of visits per month is nearing 1500 (some 60 visits per day, of which more than half from outside Romania).
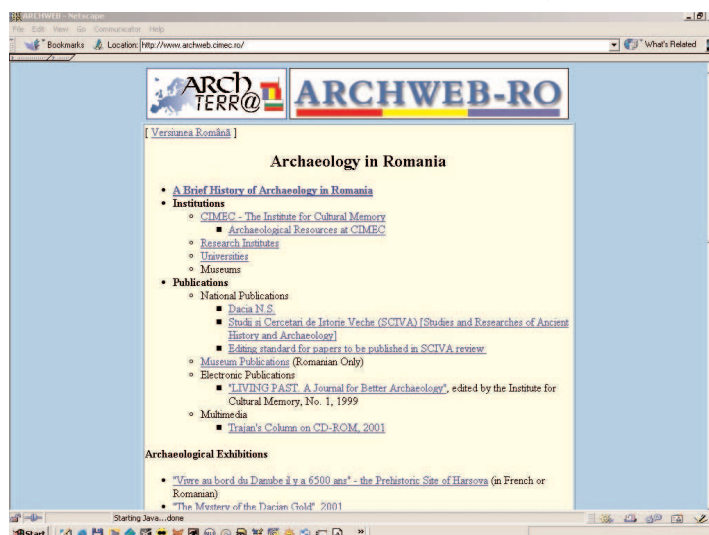
Figure 7. ArchWEb Romania

2.3.3 Institutional Web Pages

At the outset of the project, partners MAP and CIMEC already possessed fairly extensive web pages, whereas partner AIM had only a very brief web page hosted by the Bulgarian Academy of Sciences. A full set of institutional web pages was therefore developed for the latter under ArchTerra, and published on the AIM server. The web presences of MAP and CIMEC were significantly enhanced by the addition of new content, functionality, and layout, as briefly described in the sections below. These institutional web pages have an important exemplary function on the national level, demonstrating as they do the power of Internet for the dissemination of information, opening up communication channels between the public and professionals, and the use of up-to-date web interface and database applications.

Archaeological Institute with Museum (AIM)

For the home page of the Archaeological Institute with Museum in Sofia, go to http://aim.bas.bg (in Bulgarian and English). The AIM web site is divided into 7 sections, relating to Historical background & structure; Administration; Research departments; National archaeological museum; International projects; Publications; and the Museum Information System.

The content of each is detailed below.

- The first section describes the more than 100 years of history of the National Archaeological Museum and the Institute of Archaeology which were united after World War II in the present-day AIM, which is part of the Bulgarian Academy of Sciences.
- The second section lists the names of administrative personnel as well as telephone and fax numbers of all the administrative bodies of the AIM - BAS.
- The third section describes briefly the six research departments of the AIM - BAS with the number of members, the research priorities as well as some of the major excavations directed by the department members.
- The fourth section describes the collections held by the Museum under the following headings: Prehistoric; Thracian and Classic; Medieval; Numismatic; Epigraphic; and Treasures. It also contains information about the temporary and permanent expositions of the Museum.
- The fifth section lists the foreign institutions that AIM-BAS has permanent contacts with, as well as the international projects which AIM - BAS participates in.
- The sixth section lists all the periodicals and series edited and published by AIM – BAS.
- The seventh section offers two types of access to the Museum Information System: one for the general public (anonymous access), the other for registered users.

Muzeum Archeologiczne w Poznaniu (MAP)

The homepage of the Archaeological Museum of Poznan was enhanced with the addition of sections on permanent and temporary exhibitions (the former on "Prehistory of Greater Poland" and "Death and Life in Ancient Egypt", the latter on "Gods, Graves and Mummies: way to eternity in Ancient Egypt" and "Flight into the past. Aerial archaeology in Britain"), on the management of Archaeological Heritage (including a slide presentation "Air Photos in Archaeological Prospection"), and on excavations carried out by the Museum in Poland and abroad (fig. 8). Pages were also added on the Museum's seat - the Renaissance Palace of the Gorka Family (photo tour), on the archaeological staff of the Museum (including CV's), its publications (the yearbook, a series on African Archaeology, and a series on computer applications), and its other activities. User statistics for the MAP home page indicate that it is already functioning as one of the main gateways to Polish archaeology within Poland itself, with over 400 hits in January 2001. The museum's web page is at http://www.muzarp.poznan.pl/museum.htm

Figure 8. Web page of Archaeological Museum of Poznan

Institutul de Memorie Culturala (CIMEC)

Founded in 1978, CIMEC (Institute for Cultural Memory) is a public institution under the Romanian Ministry of Culture and Cults, and a national organisation for the computerised cultural heritage

record (fig. 9). CIMEC is an institutional member of ICOM (International Council of Museums) and is financed mainly by the Ministry of Culture and Cults on a contractual basis. Its main activities are:
- collects, processes, develops and disseminates information concerning movable and immovable cultural heritage, theatre performances, cultural institutions, bibliographic records and cultural events;
- maintains the national databases and other computerised cultural information resources;
- develops tools for the collections documentation (artefacts description rules, data standards, terminological thesauri, authority files, software etc.);
- is an editor of various publications, in both classical and digital formats;
- provides consulting and technical assistance for museums, libraries and other cultural institutions.
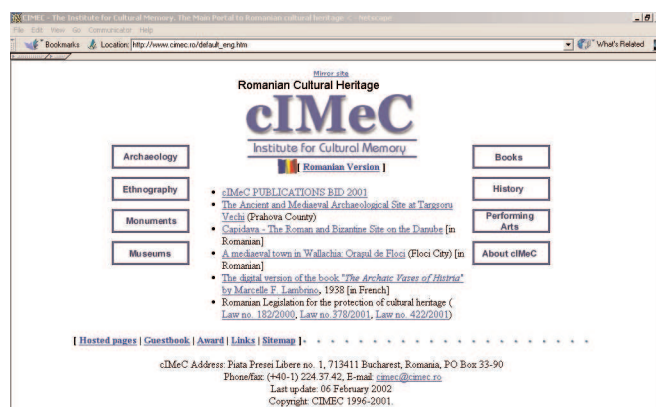

Figure 9. Web page of CIMEC

CIMEC also maintains (mainly) computerised records about the movable heritage (760,000 museum objects and 87,000 rare books), 32,500 monuments and sites that constitute the immovable heritage, some 11,000 theatre performances, 1,000 archaeological excavations, 2,400 museum specialists. In the 2 years of the ArchTerra project, the server capacity of CIMEC increased threefold (both in storage capacity and processing speed). Besides, the mirroring of the CIMEC site by project partner CILEA improves considerably the access of our foreign visitors (the quality of our communication infrastructure is still low). For the home page of the Institute for the Cultural Memory in Bucharest (Institutul de Memorie Culturala), go to http://www.cimec.ro (in Romanian and English). The (Romanian) Archweb site is now one of the most successful parts of CIMEC's website, attracting a daily average of 2,000 visitors.

2.3.4 The Joint Demo
This collaborative web site was constructed by the CEC partners in the ArchTerra Consortium in an attempt to demonstrate how the Internet may be used to bring central and eastern European archaeology to the attention of the general public, and how providers of archaeological information in those areas may benefit from sharing and exchanging information. Accordingly, this web site:
- brings to the fore the archaeology of Eastern Europe, which has until now been severely underrepresented on the Internet, as in western printed media;
- demonstrates the international character of the archaeological heritage - past cultures did not respect modern national boundaries, and mineral resources (for example) were traded and transported across large stretches of Europe.

This web site will be enhanced in the future with a demonstration that 'distributed' online exhibitions can be constructed by a collaboration between information providers (eg, archaeological museums) from several countries. An online artefact database, to be run at CILEA, will demonstrate how building blocks for an online exhibition can be assembled over the Internet. In the meantime, the 'look and feel' of the three sections supplied by the participating organisations are largely left intact so that the user remains aware of the distributed nature of the information.

Credits
This web site was produced by a collaboration of organisations active in studying and preserving the archaeological heritage in eastern Europe. From Bulgaria, the Archaeological Institute with Museum

(AIM) and the Department of Archaeology of New Bulgarian University (NBU), both of Sofia, participated. From Romania, the Centre for the Cultural Memory (CIMEC) at Bucharest, and from Poland, the Archaeological Museum of Poznan (MAP). The scenario was prepared by Dr Andrzej Prinke of MAP. The content (texts and images) were prepared by the staff of the four archaeological institutions involved, and the HTML coding as well as the user interface was performed by III, Ltd., of Sofia (Bulgaria). Dr Stefan Alexandrov and Mr Michail Vaklinov were responsible at AIM for the preparation of the materials; Prof Gatsov at NBU; Dr Irina Oberlander-Tarnoveanu and Ms Corina Bors at CIMEC; and Dr Andrzej Prinke at MAP.

2.4 Information systems for Eastern European Archaeological Museums

Eastern European archaeological museums, like most of their western counterparts, are small and poorly staffed. Often the staff is unaware of the possibility, or unable to afford, migration of its records to a digital Museum Information System. Such migration is a precondition for being able to share information about its holdings over Intranets and the Internet. One of the core objectives of the ArchTerra project has therefore been to supply a model information system for archaeological museums according to existing international standards, which could then be adapted to the needs of each individual museum and could be implemented at a relatively low cost. The current section describes the software developed by ArchTerra in order to offer CIDOC-compliant MIS to archaeological museums, and the work done to implement such an MIS at partner AIM.

2.4.1 The CIDOC Model and its Implementation
The data model developed under the project was initiated as an implementation of the CIDOC model of Information groups and Information categories. During the course of the work it was essentially enhanced in line with the CIDOC Object-oriented data model recommendations as well as existing practice at the Archaeological Museum in Sofia, so as to fully match current practice. The result is a set of three models which are fully CIDOC compliant and can be used for implementation of a broad range of museum information systems in Archaeology and other domains of cultural heritage:
- Object-oriented museum data model. This is developed in a standard subset of Universal Modeling Language (UML) and can be used for implementing either an object-relational or a standard relational database with museum information;
- Procedural data processing model. This is developed in a standard subset of a highly intuitive visual flowchart language and can be used as a technical reference specification for developing desktop, client/server and Internet/intranet information systems;
- Set covering user profile. This is developed as a list of overlapping role sets with attached functions and can be used for implementing flexible database access discipline and for database administration of different types of information systems.

From this model two different databases have been generated: an object-relational database for Oracle 8 RDBMS, and a standard relational database for Oracle/MS SQL Server RDBMS. In addition to these databases, three separate client/server development tools have been experimented with for actual implementation of the museum information system:
- Oracle Developer for implementation of both client/server and Internet/intranet information systems using Oracle Designer generated Forms/PL/SQL code;
- Oracle Jdeveloper for implementation of both client/server and Internet/intranet information systems using visually programmed Java code;
- Oracle WebDB for implementation of a simple Internet/intranet information systems using visual templates.

From these experiments it became clear that the best combination would be to use a Developer generated forms application for local client/server operation together with visually programmed Java servlet applications for remote Internet access to the museum database.

Implementation of a CIDOC-compliant Information System
The full CIDOC model was implemented as a model for a museum information system according to the recognized international standard for networked museum information systems with several clients and a centrally located server. Since the CIDOC model is generic, applicable to any type of museum, it must be adapted whenever a MIS for a specific museum is implemented. In the case of

AIM, the CIDOC model was both reduced in some respects, and extended in others; in the case of MAP, the CIDOC model had already largely been implemented, and therefore did not need to be replaced – it only had to be interfaced with the other two systems following the CIDOC standard. In addition, there appeared some legal problems relating to rights of access to certain types of information. Because of this, III has implemented a second, scaled-down version of the system, which more closely follows the existing practice at AIM; it is much lighter and can be used equally well by other small museums in the country.

The full CIDOC-compliant information system was implemented as a server-side intranet/internet application using Oracle WebDB. It behaves as a Web site, generated dynamically from database content and controlled from a single point. It is administered using series of Web templates for site, user and content management.

Museum Site Template

The Museum Site template developed under ArchTerra consists of two parts. Both components of the template are consistent with the CIDOC Model of information groups and categories, and can be used for implementing museum information systems for heritage institutions across Europe. Both templates have been developed using CASE methodology and tools from Oracle, Corp.

Museum Database. This part of the template is entirely independent from the database in use, since it relies on features found in any standard relational databases, including Oracle, Informix, Microsoft, Sybase and public domain databases like mySQL. Initially this part was developed as an object-relational database after the Object-oriented enhancement of CIDOC, but for efficiency and universality it was also developed as an entirely relational database. Three different databases have been generated on the basis of this template under the project – one relational database compatible with Oracle 7 Server, one object-relational database compatible with Oracle 8 Server and one relational database for Microsoft SQL Server.

Museum Information System. This template, although designed using specific Oracle client modules for interaction with the database (Forms and PL/SQL), is relatively universal; it can be directly implemented for other clients, such as Microsoft Visual Basic, and adapted for any public domain tools such as Java.

Scaled-down Implementations

Characteristically, the main potential users of museum information systems under ArchTerra – TAIM, MAP, and CIMEC – had widely varying work practises and software environments – thus, Oracle was used in AIM, VisualFoxPro in Poznan and MS SQL_Server in Bucharest. None of them saw any advantage in relinquishing their current practice for the full CIDOC standard; they therefore implemented their own versions of museum information systems as follows:

- AIM: Based on the experience gained during development of the CIDOC-compliant information system, III has implemented scaled-down version of it after the recommendations of AIM and using the same technologies – Oracle as a backend database server and WebDB as a dynamic site for approaching the museum database over the intranet/Internet. See Deliverable 11 for the AIM IS installation manual and user guide.

- MAP: A new, upgraded version of the existing database system MuzArP, based on VisualFoxPro with SQL data retrieval, has been developed. This information system combines data on archaeological sites, their study (excavations), and results (museum objects as well as multi-aspect information on them), covering all basic needs of the staff of an archaeological museum: research, inventory, management of collections and of sites & monuments. It follows the complex data model (351 data fields) developed and tested in practice in the earlier DOS version of the system.

- CIMEC: Together with CILEA, CIMEC implemented, tested and customised the data models and database applications built by III. Furthermore, CIMEC has developed (and is currently refining) a UML model for the following categories: places, sites, finds; terminology, topics, subject headings; time, events, entities having time spans; agents: people and organisations/groups; and bibliographic entities (conforming to the IFLA Functional Requirements for Bibliographic Records). Based on this UML model an MS-Access database schema was developed and data acquisition and browsing applications of the model were programmed.

2.4.2 Controlled terms for indexing and retrieval

No MIS could function without proper controls on the information that is being entered into it. One

of the main tasks has therefore been to establish new, or convert existing, controlled term lists (vocabularies, glossaries) for use in the AIM IS. For AIM, the controlled term lists were established according to CIDOC standards, resulting in 21 lists of terms which describe material, technical, scientific and administrative aspects of the museum's holdings (see above). Besides being directly incorporated into the AIM information system, the terms have served as a 'seed' for the building of a multilingual glossary of archaeological terms.

2.4.3 Digitisation of holdings
It was originally intended that several thousand of AIM's holdings would be fully digitised during the course of the project. In the event, the start of this work at AIM was much delayed by practical problems of access to the equipment installed under ArchTerra, and even to the data itself. The internal administrative and security procedures at AIM made it impossible to set up efficient digitisation procedures. Notwithstanding this, by the end of the contract period almost all the prehistoric holdings of AIM have been digitised, and the first 100 of these enhanced with digital images.

2.4.4 Evaluation
The Consortium's experience with the implementation of MIS according to the CIDOC Model highlights the many practical difficulties in promoting state-of-the-art IT solutions, not just in eastern European archaeological museums, but in small museums anywhere in Europe. These difficulties appear to be caused by two separate sets of factors. Firstly, the climate of (sometimes extreme) financial hardship under which archaeological institutions in large parts of Europe must operate, translates itself both directly in a lack of (IT and financial) resources needed, and indirectly in inadequately trained personnel and conservative administrative procedures. Whereas the ArchTerra project successfully addressed the former, the latter result in an institutional climate of resistance to innovative IT approaches to Museum Information Systems.

The second set of factors is related to the basic question of why should museums adopt the CIDOC Model. Undoubtedly the replacement of an existing MIS by a new one is a much more difficult process than the initial adoption of an MIS would be. On the one hand the work involved in data migration, changing work practise, and re-training staff, plus the fact that software revisions and bug fixing are no longer handled in-house, could justifiably deter museum staff from supporting such a change-over; on the other, the advantages of the CIDOC Model over the existing data model might not be apparent or sufficiently large to warrant it. Considerations of international best practise may appear irrelevant to most museums, which are merely looking for an MIS that will make their existing practise more efficient and effective.

A more general set of problems that became apparent in the course of the work reported here, was the unfamiliarity of most professional archaeologists and museum staff with modern concepts of information structuring. Thus, ArchTerra consortium partners had to spend more time than expected in first learning about, then explaining to others, such basics as the difference between glossaries, restricted term lists, and thesauri. It must be noted that even if professionals *are* aware of proper methods for information structuring, they often do not follow these methods; for example, it was found that the set of archaeological terms in the Dutch National Central Catalogue system (NCC), constructed in the 1990s by subject specialists, was severely polluted with illogical and missing entries.

The administrative barriers to data entry imposed at AIM, added to the delays implementing a scaled-down MIS, have led to disappointing results in the development of digital content. Although AIM staff is now in a position to continue digitisation of all its holdings, at the end of the project the work had not progressed sufficiently to allow an evaluation of the software interfaces and digital records to take place.

The experience gained on these issues by the project partners, if properly disseminated, will come to good stead in future attempts to modernise information systems in small museums across Europe.

2.5 Research

The ArchTerra Consortium conducted research in the area of knowledge organising systems (NKOS), specifically into the development of structured term lists, their use as thesauri in user

interfaces, and their translation into multiple European languages and scripts. Two prototype NKOS were built; one forms part of the MIS discussed in section 2.4, the other is presented in abridged version in section 2.5.1 (the full version is appended as Deliverable 12). Three sets of term lists, produced to explore different aspects of international and multilingual indexing and retrieval of archaeological information, are briefly introduced in section 2.5.2 (the documents themselves are appended = deliverable 13). Together, these documents provide the materials needed to present the potential benefits and problems of European archaeological thesauri to a wider professional public, to open up discussion, and to prepare proposals for follow-up projects.

2.5.1 Indexing and Retrieving Archaeological Resources on the Internet

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.

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.

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 multiliguicity, 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.

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 – multicultural – and conflicting meanings of the same term. Among the practical approaches to consider are:
- taking an arbitrary decision, e.g. to abide by the definitions given by some authority, and
- ensuring professionalism in the translators (avoiding debates about the quality of a translation).

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. 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 standard, 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.

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-Apennine' 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).

Relation Types

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

1. *Class* (CL; top-level descriptive category to which the term belongs, e.g. 'Chronological Period')

2. *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.)
3. *Broad Term* (BT; indicates the generic term of which the current term is a specific; a.k.a. 'is a kind of')
4. *Narrow Term* (NT; indicates the specific terms of which the current term is the generic; a.k.a. 'contains')
5. *Related term* (RT; indicates all relation types not covered by BT and NT)
6. *Used for* (UF; indicates any alternative but non-preferred terms with the same meaning; a.k.a. 'is a preferred term for')
7. *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 it 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.

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. The results of the primary search will be accompanied by dynamically generated options for widening or narrowing searches which constitute the secondary retrieval system.

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 $\leftrightarrow$ 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.

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 review both these issues, using examples from ARGE's prototype thesaurus maintenance interface.

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. An option to call up the helper window containing the current thesaurus terms and descriptions was also added.

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 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.

2.5.2 Multilingual Dictionaries and Thesauri for European Archaeology

The ArchTerra project produced three sets of terms in order to explore different aspects of international and multilingual indexing and retrieval of archaeological information:

1. the thesaurus of indexing terms implemented within the Museum Information System of the Archaeological Institute with Museum of Sofia (AIM; English version; the implementation itself is bilingual Bulgarian-English);
2. a multilingual dictionary of archaeological terms, covering periods and cultures in the archaeology of Central Europe. The five languages used by the dictionary are: Bulgarian, Polish, Romanian, Italian, and English;
3. the thesaurus of indexing terms implemented within ARGE Virtual Library Management System 3.0 (English version; the implementation itself is language independent).

Each of these explores different aspects of the international and multilingual indexing and retrieval of archaeological information. Document A, produced by staff at AIM, exemplifies the specialist use of exhaustive restricted hierarchical terms lists (in this case for the classification of the holdings of AIM according to the classes *materials*, *chronology*, and *subjects*). Document B, the result of a collaboration between several Consortium partners co-ordinated by CIMEC, concentrates on problems in the translation and multinational / multicultural use of archaeological thesauri in a first attempt to define a set of terms valid across Central Europe and find equivalents in several important languages. Document C, produced by RUG, represents a more generalist, 'opportunistic', approach to the construction of archaeological thesauri by attempting to organise indexing terms of the Internet resources listed in ARGE, and explores in particular problems relating to the process of construction itself.

2.5.3 Evaluation

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 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 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).

Multicultural translation. Although databases may be structured to be language-independent, and web interfaces to those databases may be made available in multiple translations and transliterations, and the technical problem of mixing and correctly displaying two or more *scripts* (eg, Latin and Cyrillic) in a single frame within a web page is largely solved, coding multilingual Internet applications remains far from trivial. Even more intractible than the immediate and practical problems of translation, the challenge in constructing generic European archaeological thesauri is in dealing with multi-culturality and shifting meanings of terms; these may require thesauri to be abstracted into semantic structures rather than relying on a single 'base' term set and language.

2.6 Dissemination and Exploitation

In view of the unfamiliarity of most professional archaeologists and staff of archaeological museums with the issues confronting the ArchTerra Consortium, the dissemination of its aims, objectives, approaches and results is an important part of the project's work. During the project's run-time, active dissemination has mainly been done through a series of papers presented at various international symposia, many of which have already been published or will be published in the near future (section 2.6.1). As passive means of dissemination, an extensive Internet presence has been established for the project (section 2.6.2). However, the aims and the objectives of the Consortium also require that much of the work must remain 'in progress', and will be maintained and developed

further both by the current Consortium partners and, after dissemination, by others. For this reason a detailed exploitation plan has been established (section 2.6.3).

2.6.1 Presentations and Publications
Presentations
The Consortium partners presented posters and distributed leaflets during the annual international conferences of the two most relevant organisations of professional archaeologists: the European Association of Archaeologists (EAA), established in 1993, is the main international professional organisation, currently with almost 2000 members; the Computer Applications and Quantitative Methods in Archaeology (CAA) annual meetings provide the primary venue for European researchers involved in archaeological computing.

- At the *1999 CAA Annual Meeting* (Dublin, IE), Dr A Prinke gave a paper "Extending the European Archaeology Web over Bulgaria, Romania, and Poland", presenting the ArchTerra project. The paper was subsequently revised and will appear in the proceedings of the CAA 2000 conference.
- At the *5th EAA Annual Meeting* (Bournemouth, UK, September 99) two papers were presented by partners from RUG and CIMEC.
- At the *1999 Digital Resources in the Humanities Conference* (London, UK, September 99), a paper was presented by Mr Van Leusen on the subject of "Serving Archaeology: the ARGE Virtual Library management system".
- At the annual *Meeting of the UK chapter of CAA* (Durham, UK, February 2000), Dr Vassilev gave a paper on 'Reducing the complexity of CIDOC OO model through ontological minimisation: the ArchTerra experience'.
- At the *2000 Annual Meeting of CAA/UISPP* (Ljubljana, SI), a poster and leaflets presenting the three national ArchWebs of Romania, Bulgaria, and Poland was displayed, and a workshop on approaches to multilingual indexing and retrieval of archaeological information was held. A paper presenting these approaches will be published in the conference proceedings.
- At the June, 2000, Raphael Programme conference *Mapping the Future of the Past. New Information Technologies for Managing the European Archaeological Heritage* (Seville, E), A. Prinke presented a paper on 'The activities of Poznan Archaeological Museum in introducing information technology to archaeological heritage management (1986-2000)'
- At the *6th EAA Annual Meeting* (Lisbon, PT, September 2000), the project was represented with multiple posters, leaflets, papers, and a workshop (deliverable 18). A poster on the ArchTerra project was presented; leaflets on ArchTerra and Mediolanvm were handed out. Dr Vassilev and others presented a paper on 'Drill-down navigation inside archaeological museum databases'. In the workshop on ArchWebs, project members presented papers on "East European Archaeology at your fingertips: Archaeological Web Nodes for Bulgaria, Poland and Romania", "ArchWeb-PL: Gateway to Polish Archaeology", "ArchWeb-BG: Gateway to Bulgarian Archaeology", and "LASE, MEDIOLANVM, ArchAgenda: Central services for national ArchWebs". Abstracts of all papers were published in the pre-conference handbook of the 6th EAA.

The ArchTerra project was also presented by project member CIMEC in a radio broadcast on Radio "Romania Cultural", with national coverage), and by MAP during the Euroconference "VAST – Virtual Archaeology between Scientific Research and Territorial Marketing" (Arezzo, I), the NATO Advanced Research Workshop in Leszno (Poland), and the International Coloque on computer applications in archaeology at European University Viadrina (Subice – Frankfurt a/Oder), all in November 2000.

Publications
Articles on several aspects of the project appeared in ICOM-News (the newsletter of the International Council of Museums), and in two separate issues of the bulletin published and distributed widely within Italy by CILEA (see following section); scientific articles will appear beginning in early 2001 with the publication of the proceedings of the CAA2000 conference.

NEGRONI CATACCHIO, N., GUIDETTI, L., CAMNASIO, M., FERRARI, R., VAN LEUSEN, P.M., 2000, MEDIOLANVM European Fieldwork Server (CAA Conference – April 18-21, 2000 Ljubljana, Slovenia) *Bollettino del CILEA* 73 (Giugno 2000): 7-12.

OBERLÄNDER-TârNOVEANU, I., 2000, ArchTerra: Sharing Archaeological Knowledge across Europe, *ICOM News. Newletter of the International Council of Museums*, volume 53(3):7.

VAN LEUSEN, P.M., 2001, Indexing and Retrieving Archaeological Resources on the Internet: A prototype multilingual thesaurus application, in Z. STANCIC, T VELIANOVSKI (eds.) *Computing Archaeology for Understanding the Past, CAA 2000* (BAR International Series 931) Oxford: Archaeopress, 303-8.

VAN LEUSEN, P.M, PRINKE, A., 2001, ArchTerra. Extending the European Archaeology Web over Bulgaria, Romania and Poland, in Z. STANCIC, T VELIANOVSKI (eds.) *Computing Archaeology for Understanding the Past, CAA 2000* (BAR International Series 931) Oxford: Archaeopress, 323-325.

VASSILEV, V., STOIEV, I., GAYDARSKA, B., ALEXANDROV, S., NEHRIZOV, G., VAKLINOV, M., 1999 Museum Information Systems: CIDOC data model implementation in the ArchTerra project *Bollettino del CILEA* 69 (Settembre 1999): 15-30.

2.6.2 Internet presence

ArchTerra's Internet presence makes itself felt at several levels: internationally, nationally, and at tle level of individual institutions. The project itself of course has a home page as well, giving access to all the resources developed under it.

Project web pages and services

The ArchTerra home page, for general access to and information about project resources, is available at archterra.cilea.it. To increase its chances of being found by interested Internet users, this URL was submitted to several international search engines (eg, Yahoo!, Google, DMOZ) in July 1999, and missing project details were added to the CORDIS database maintained by the EU in August 2000.

The central internet service for field work opportunities MEDIOLANVM was officially opened on June 1st, 2000, for general use, at archterra.cilea.it/ mediolanum/. The ARGE Virtual Library, for general use, remained accessible throughout the project at www.let.rug.nl/arge/. Finally, the demonstration pages for 'distributed' archaeological exhibitions are available at archterra.cilea.it/ exhibits/demo1/.

National hosts and services (ArchWebs)

Of the three ArchWebs developed under ArchTerra, two (for Romania and Poland) have been accessible (barring network problems) from the start of the project while the third (for Bulgaria) came online only as the project ended. To improve the accessibility of these web pages, CILEA mirrors the ArchWeb sites at http://archterra.cilea.it. The Romanian ArchWeb node will be found at www.archweb.cimec.ro; the Polish ArchWeb node at www.muzarp.poznan.pl/ archweb/; and the Bulgarian ArchWeb node at archweb.nbu.bg.

Partner hosts and services

In the course of the ArchTerra project, the institutional web pages for each of the CEC partners were significantly enhanced in order to provide exemplars for archaeological instutions across Europe.

For the Archaeological Museum of Poznan (Muzeum Archeologiczne w Poznaniu), go to http://www.muzarp.poznan.pl/muzeum/ (in Polish and English); for the home page of the Centre for the Cultural Memory in Bucharest (Centrul de Memorie Culturala), go to http://www.cimec.ro (in Romanian); for the home page of the Archaeological Institute with Museum in Sofia, go to http://aim.bas.bg (in English).

2.6.3 Exploitation plan

The ArchTerra Consortium consists of partners that have no overriding commercial interest in the exploitation of project deliverables. Instead, they are more concerned to allow unhindered distribution of prototypes, tools and guides to other archaeological user communities on a 'shareware' basis (that is, on the basis of distribution and maintenance cost recovery). The goals of the ArchTerra project call for a continued activity of the Consortium partners after the formal end of the project. The partners' intentions to continue the work and exploit the results of the project may be grouped under the headings of maintenance of services, creation of additional content, further dissemination activities, and further development and enhancement of software products. These themes are set out in detail in the exploitation plan below.

Maintenance of services

Under 'maintenance' we here understand both the upkeep of the service itself, and the continued updating of its contents.

- RUG and CILEA will maintain, respectively, the ARGE and MEDIOLANVM central services for a period of at least 3 years from the end of the contract;
- CIMEC, MAP, and NBU will maintain, respectively, the national ArchWeb sites of Romania, Poland, and Bulgaria;
- CIMEC, MAP, and AIM will maintain, respectively, their database information services;
- CILEA will continue to host and maintain the ArchTerra project web site and mirror sites of the three national ArchWebs of Romania, Poland, and Bulgaria for a period of at least 3 years from the end of the contract.

Creation of additional content

- AIM – will continue the digitisation of its holdings within its MIS; it will add the following new sections to ArchWeb-BG: Archaeological sites excavated by AIM-BAS members, Contents of the AIM-BAS editions, Exhibition projects and descriptions, and New archaeological and museum projects;
- MAP – will extend its web pages with the full contents of the Poznan Museum's yearbooks;
- CIMEC, MAP, AIM – will use the "know-how" gained within the ArchTerra project for realising further virtual on-line exhibitions;

Other related activities include the publication in 2002 of a multimedia CD-ROM on Trajan's column as a result of the collaboration between partners CIMEC CILEA and III.

Dissemination activities

All partners will remain active in the promotion of access to Internet resources on European archaeology. Specifically, they will provide practical aid in setting up (and, where necessary, hosting) new ArchWeb nodes in other countries of Europe; and they will continue to disseminate leaflets on project deliverables.

CIMEC – will organise workshops on ArchWeb-RO at the Faculty of History of the University of Bucharest (special seminary for Methods in Archaeology) and at the Annual Meeting of Romanian Archaeologists (mid-summer 2001). It will also continue to provide its museum information system and conversion software to Romanian museums at low cost;

- MAP – will organise the XIIIth Report Conference on Archaeological Research in Greater Poland in 1998-2000 (April, 2001), during which the ArchTerra project will be the subject of the keynote speech, and the MAP and ArchWeb-PL web pages will be presented;
- CILEA – will publish papers on the ArchTerra project and the Mediolanum service in Italian language scientific journals;
- AIM – will organise a presentation of its MIS and the ArchWeb-BG web pages at a meeting of archaeologists from the regional archaeological museums of Bulgaria on the premises of AIM (February 2001);
- NBU – will organise a presentation of the ArchWeb-BG web pages during a special workshop to be organised by the Department of Archaeology of the NBU in March 2001;
- III – will offer its implementation of the CIDOC data model developed under ArchTerra to CIDOC for the purposes of dissemination and further development.

Development and enhancement of software products

The partners will continue to work towards a joint ODBC database system for the storage of, and access to, national archaeological records (co-ordinated by CIMEC); and they will take part in a working party to plan for the establishment of an internet gateway to European Archaeology under the aegis of the European Association of Archaeologists (co-ordinated by RUG).

- CILEA – will, in collaboration with the University of Milan and an Italian museum yet to be chosen, test implement another instance of the museum information system, in Italian.
- III – the CIDOC based MIS will be proposed to all museums in Bulgaria which are willing to implement/adapt it for their own use; a national information system fort the most important objects and documents in Bulgarian museums and archives is already under development by III and the National Chamber of Museums in Bulgaria. It is based on the CIDOC compliant MIS developed by Dr. Vassilev;
- RUG, CIMEC – will continue development of multilingual thesauri and glossaries, and disseminate these through ARGE.
- NBU, MAP, CIMEC, CILEA – III will develop a MySQL/PHP-based directory application for inclusion in the national ArchWebs, which will contain details of people and organisations related to archaeology; the contents of the three will be made available as a merged directory application through the ArchTerra web site at CILEA.