

An Integrated Information System for Archaeological Data Management: Latest Developments

Vittorio Fronza, Alessandra Nardini and Marco Valenti

LIAAM (Laboratorio di Informatica Applicata all'Archeologia Medievale)
Dipartimento di Archeologia e Storia delle Arti – Università degli Studi di Siena
via Roma 56, Siena 53100, Italy
{fronza, nardini, valenmar}@unisi.it

Abstract. At other CAA meetings we have already introduced the main experiences of our laboratory. It is our intention, here, to discuss the latest developments of our archaeological management system, called OpenArcheo. We work on different scales (from regional surveys down to detailed records of all the finds). In particular all the data is administered within a data model made up of five components: different GIS platforms, an alphanumeric database, a media database, a spreadsheet with Harris's matrixes, multimedia presentations; these are linked by a system level application. OpenArcheo is constantly updated in order to fulfill new functions and manage new in-depth data; basic concept of our system are the multidirectional links between information types, which allow the user to retrieve and process all the information related to a specific archaeological feature.

Key words: Approach to Computer Applications, Intrasite GIS, Integrated information system, Data Architecture

1 Introduction

Application of information technology to archaeological research started, at the Medieval Archaeology Area of our Department, during the late '80's-early 90's; the process has been led, from the beginning on, by Riccardo Francovich and Marco Valenti (regarding our history and guidelines, the principles which inspired our information system, its data architecture and its functionalities refer to Francovich and Valenti 2000; Valenti 2000). Our activity has always aimed at developing products suitable to manage and analyze archaeological data. The growing interest of the archaeological community and the encouraging results in this field determined, in 1998, the set up of a specific laboratory, called LIAAM (Laboratory of Information Technology Applied to Medieval Archaeology).

Today the LIAAM is made up of more than 40 people, comprehending many researchers and about 20 students with yearly scholarships; we manage 15 open area excavations, 2 provincial landscape survey projects, 2 complex urban archaeology projects. Our activity tries to cover most of the fields involving digital tools applied to archaeology, with heavy concentration on intrasite and landscape GIS, database management of all archaeological evidence, spatial analysis, multimedia production and web publishing.

1.1 A few statements

Our particular approach to the use of computers in research projects can be resumed in a few points.

Archaeologists and not Computer Scientists. The statement can be interpreted in two different ways, both of which apply to our experience.

In first place, we believe that archaeologists should manage personally every step of the digital management, from acquiring

and processing data to scientific publication and wide-public presentations. On the other hand it happens too often that, in the process of elaborating new research methodologies related to digital management, archaeologists acquire an absolutely high-end knowledge of particular computer science matters; the risk is to be being caught up in a research field which has its own problems and methodologies, having little to do with archaeology. A computer scientist pursues formal perfection of a digital application, efficiency in terms of hardware and software architecture; in the research of a solution he privileges technical aspects, which allow taking advantage of all the resources of a digital environment. An archaeologist puts on the foreground his particular historical problem, his research methodologies, the answers he wants to get out of a data set; often he prefers an application which is less optimized under a computer science perspective, but more adherent and coherent to his research activities.

If we don't want computer applications in archaeology to remain an elite discipline we need to have a pragmatic approach to technology, in order to find an *archaeological way to computer science and not a computer science way to archaeology*. Finding a right compromise in acquiring knowledge about digital tools should represent one of the major points of reflection and discussion within the large area of computer applications in archaeology.

This first considerations influence heavily all the statements that follow.

Open Data Architecture. The vast (and often interdisciplinary) subjects covered by archaeological research, the constant improvement and enhancements of methodology and the particularities of each investigated context, make it impossible to completely master from start point every possible variable and data set a project could need to manage. An open data architecture, which allows easy and progressive integration of different information and processing systems, represents a basic concept in the development of digital tools and solutions.

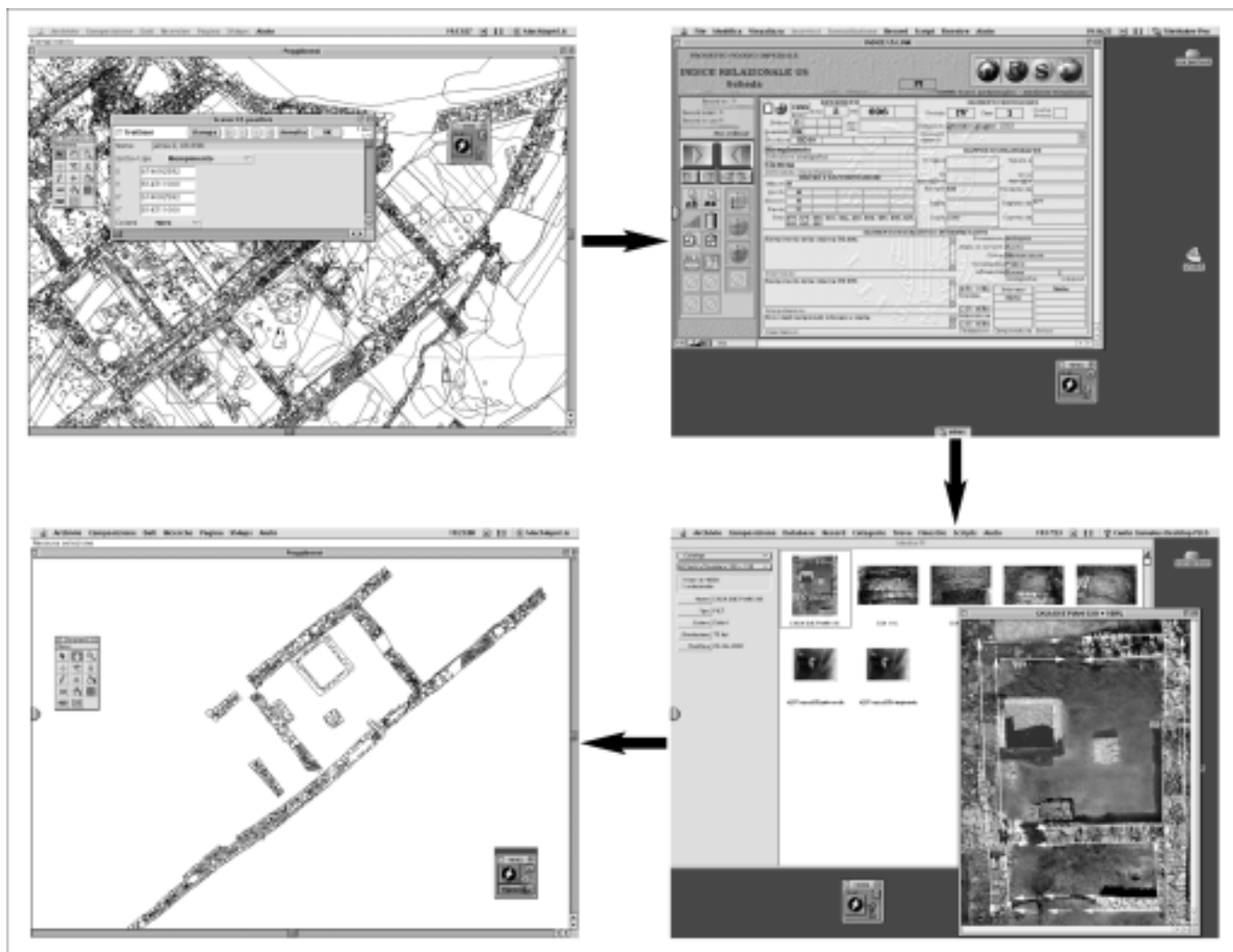


Fig. 1. An example of OpenArcheo at work on the excavation of Poggio Imperiale. In the first step a context (the fill of a cistern) is selected on the GIS; the second step shows a link to the record sheet in the database, while on the third step a link has been carried out towards the multimedia database, where one of the pictures has been opened. The last step shows a link from the selected picture back to the GIS platform.

Setting the standards: an approximation process.

Despite the considerations made about the previous statement, we believe in the possibility of creating solutions which apply to different archaeological contexts; in other words, easily exportable tools.

This can't be done in one single and comprehensive planning step: we have to get there by approximation; every time a particular project needs to store new data classes or new tools to process data, these should be integrated in an open architecture. Fitting new features within an existing data model ensures compatibility with all the previously implemented work; it means making them available to everyone and not only to the particular project they were meant for. This method can be seen as a way of setting standards in the digital management of archaeological data.

Obviously such a process is not accomplished within a short period of time; it has taken us years to reach the actual point of development, and our information system is constantly growing depending on contingent needs of the research projects managed by it.

1.2 The concept of OpenArcheo

OpenArcheo (fig. 1) has been presented and published in several occasions since 1996 (see especially Valenti 1998a; Valenti 1998b; Valenti 2000). It can be intended as our particular approach to the management of all archaeological evidence, from the single pottery fragment to comprehensive multimedia products, within an integrated and interrelated data architecture; on the basis of the statements discussed above, in such an environment the need of storing new data classes means simply adding and relating new modules or entities to the model. What OpenArcheo provides are easy-to-use tools to perform the main functions of archaeological data collection and analysis (a simple interface is one of our main concerns since the system is being used by a number of archaeologists who are low-profile users of digital technology).

2 Recent developments of OpenArcheo

Recent developments involve especially two main parts of OpenArcheo.

2.1 The GIS solution for intrasite analysis

Regarding intra-site GIS, the first years of activity at the LIAAM have focused on the creation of a codified system, open and therefore functional to store archaeological data on a GIS platform; in particular we concentrated on the design of a data model allowing efficient management of the huge mass of information produced by an archaeological excavation (Nardini 2000; Nardini 2001b; Valenti 1998).

Accomplishing this task has meant being able to use complete, complex and exhaustive databases, allowing us to store the archaeological information in its entirety; this makes it possible to easily produce real-time thematic maps, more or less complex queries and visualizations, high quality print layouts, etc.

During a second phase of experimentation we started a more in-depth use of the GIS tool, aiming at the production of new information by employing analytical functions; we engaged in the creation of distribution maps regarding archaeological finds, by applying quantitative and spatial methods; another field of application has been the processing of data in order to obtain predictive models of excavation frameworks.

One of the first applications of this kind has been the creation of a GIS platform for cemeterial contexts, testing it on the early medieval necropolis in Poggibonsi, (115 burials of Carolingian age; Francovich, Nardini, Valenti 2000). The goal was to make a better use of the interesting paleopathological and anthropological data produced by the excavation, trying to apply spatial analysis and mathematical-statistical tools to such a particular subject. In this platform it has been necessary to shift from the stratigraphical context to the single bone detail level; every bone has been classified on anatomical and pathological basis (anatomical part, pathology, level of pathology, probable cause of disease etc.), derived from anthropological analysis. Another entity involved has obviously been the skeleton, which has been defined by sex, age of death, burial orientation etc. In this way it has been easy to obtain distribution maps, for example, of sex, age of death, pathologies, etc. At the same time the system allowed us to put complex queries and immediately translate the obtained data in statistical terms of incidence of different conditions on the totality of the skeleton sample.

Spatial and statistical analysis has been also applied in order to build predictive models of intrasite settlement patterns. An example returning satisfactory outcomes has been carried out on the low medieval small-town located on the hill of Poggio Imperiale in Poggibonsi (Siena), where we have been digging since 1993 (2 hectares have been excavated on a total extension of about 12 hectares; regarding the excavation refer to Valenti 1996; Valenti 1999); the use of the above mentioned tools allowed us to obtain a hypothetical extension of the site during the second half of 12th century, when the urban plan was based on modular buildings (first results are presented in Nardini 2000; excavations during the 2001 campaign have confirmed the predictive model).

One last interesting case of analysis is related to the archaeological remains (refer to Boscato, Fronza and Salvadori 2000; Nardini and Salvadori 2000). In this case the detail level shifts from layers to the set of bone fragments pertaining to the same layer (or structure, or period, etc.); at the same time the underlying database (strictly dialoguing with the GIS platform) becomes the animal bones module of our architecture. Data frequency (and eventually other statistical analysis) are fulfilled on the DBMS and then imported within the GIS platform as points related to the zooarchaeological sample; the positioning is done automatically through a geocoding file based on excavation contexts. Our aim has been to display the bones deposit, by using stratigraphical and zooarchaeological keywords in order to understand how anthropological, animal and natural factors transformed the original animal population into a fossilized sample. Experimenting GIS technology on a large open area excavation like Poggio Imperiale, has allowed us to produce important information used in elaborating diachronical models of social and economical structures; it also supplies important information to predict the archaeological potential of non-excavated areas.

2.2 The multidirectional links application

As we have already seen the basic features of our system, written for Apple Macintosh platforms, are the multidirectional links between information types (which allow the user to query, retrieve and manipulate all the data classes related to a feature). In particular, all the data is administered within an entity-relationship model made up of five components: different GIS platforms, an alphanumeric database, a media database, a spreadsheet with Harris's matrixes, multimedia presentations; these are linked by a system level application directly engineered and developed at our laboratory (for an in-depth technical description of the system see Fronza 2001; Fronza, Nardini, Salzotti and Valenti 2000; Valenti 1998a; Valenti 2000;).

Lately, some new features have been added to the system.

Preferences management (fig. 2). This functionality allows the user to easily set his own preferences, regarding mainly the link keywords. For each application it is possible to choose the field or data category acting as link keyword; standard keywords are the single excavated context identifier for an excavation or, in the case of a landscape project, the site identifier.



Fig. 2. OpenArcheo preferences; the screenshot shows the preferences section regarding the alphanumeric database.

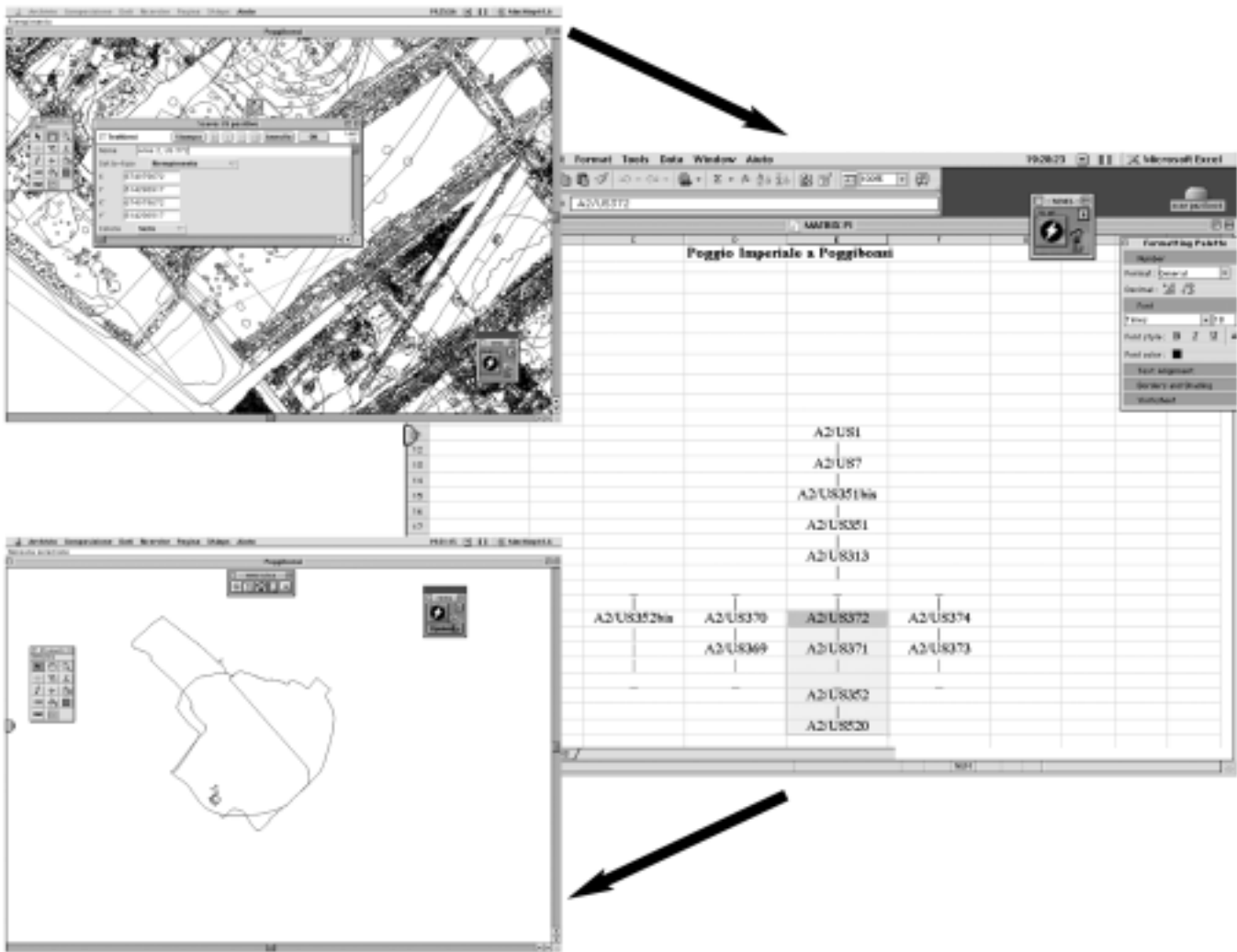


Fig. 3. Link example between a GIS object and the Harris's matrix. In the first step a posthole is selected on the GIS; the second step shows a link to the matrix (A2/US 372 is highlighted). The last step shows selection of 4 contexts on the matrix (cut and fill of the previous posthole, the ground the posthole is cut on and the layer which covers it) and a direct link to the GIS.

In many cases it might be useful to have other keywords (such as chronological phases, excavated areas, excavated structures, finds related identifiers, etc.), to put different queries to the system; for example, using a chronological phase keyword, from a selected object on the GIS we could link directly to all the pictures in the multimedia database related to the same period.

Harris's matrix management (fig. 3). Integration of Harris's Matrixes represents an important feature, which had not been provided in previous versions of OpenArcheo. Practical reasons have suggested an implementation on a spreadsheet: even though the results are graphically less accurate than those obtainable by other products (mainly organizers or vector graphics software), the cell structure adheres very well to graphical matrix representation. Another solution could have been to write our own Harris's Matrix software, implying a long development process; in fact, the already existing applications could not be used for our purposes since it is difficult to link them in our system. Moreover, we believe that archaeologists should produce their own matrixes following mainly a manual process, even though taking advantage of features related to digital recording (for example we have complex and automated controls on our database in order to avoid inconsistencies in relationships between contexts). Having added Harris's matrix management to

our system means being able to perform, for each context, direct links to and from alphanumeric records, pictures, movies, GIS objects.

Distribution analysis (fig. 4). Another new potentiality of the OpenArcheo system allows users to translate in real time statistical analysis (for example frequency analysis) processed on the archaeological finds databases in distribution maps on the GIS (for example finds concentration maps). Obviously the processed data derived from the database, can be represented using all graphical solutions a GIS application normally offers: by graduated colours based on concentration of different finds classes (or in any other way depending on query conditions), by sized symbols proportionally based on incidence level, by charts, etc. In our system, depending on the chosen detail level, every information is related to the excavated contexts or structures.

Multimedia products data class. The integration of multimedia products had been provided already in the first version of OpenArcheo, developed in 1996; in the actual configuration of the system, created in 1999, support for this kind of data had not been provided. Unidirectional links to multimedia presentations are now supported from all applications towards standalone multimedia products.

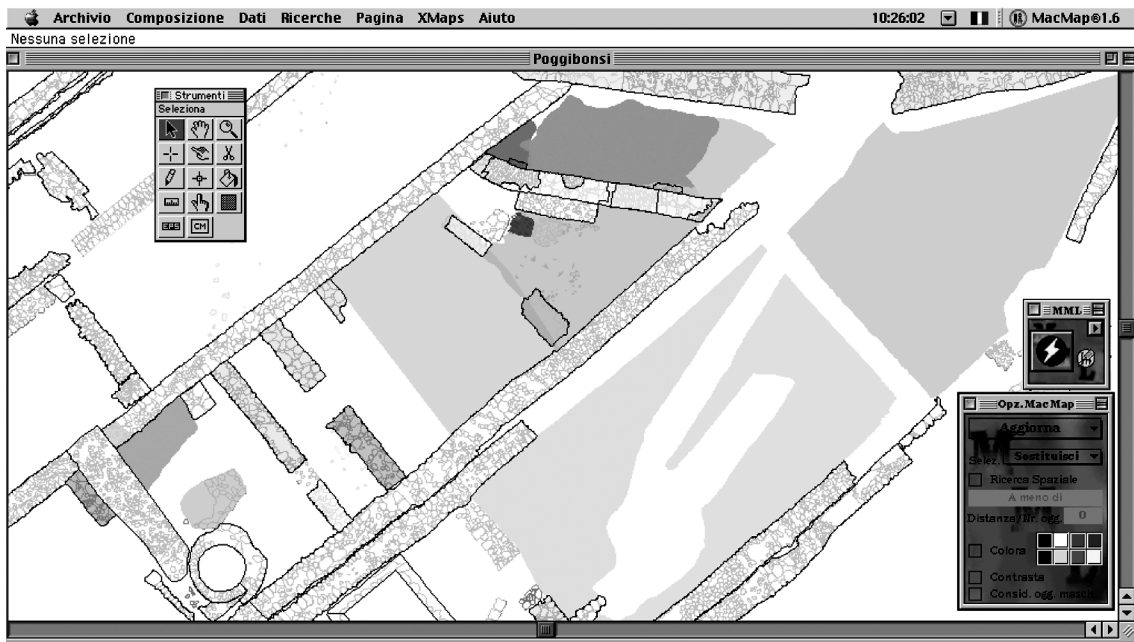


Fig. 4. The screenshot shows a distribution map obtained by using OpenArcheo. In particular it regards the concentration of kitchen pottery in Area 2 of Poggio Imperiale in Poggibonsi during the period VI/3 (first half of the 13th century); visualization is in greyscale graduated colours, from absence (white) to 27 fragments/m² (black).

It has been technically accomplished by organizing data in a simple system of folders named after the identifiers of the system and containing aliases to the related presentations; this allows us to easily produce automated queries from our user interface.

Spatial queries. In links moving from other data classes to GIS objects it is now possible to add the most common spatial research options (objects within a given distance, the *n* closest/farther objects, inside/outside an object, centred in respect to an object, etc.). For example, if we are looking at the pictures of an excavated house on the multimedia database, this feature allows us to obtain automatically on the intrasite GIS platform all the walls having the same characteristics within a distance of 50 m from the structure we were looking at; or, if we are studying a particular pottery class found on a site, it is possible to visualize immediately on a landscape GIS platform the 10 closest sites where the same class has been found.

3 In the future

Coherently with the statements discussed at the beginning of this paper, we are planning a number of major and minor improvements for our system.

A data class we expect to link to our system in the near future are web sites. More than 12.000 archaeology-related links have been collected and are already available on our site (<http://medievalarchaeology.unisi.it/NewPages/LINK/MOTO.html>). We are setting up a search engine to retrieve archaeological information on the web; this feature should be available within this autumn (actually we are producing detailed records of all links). The next step will be to integrate it within OpenArcheo; obviously this isn't related strictly to intrasite management, but applies more closely to the landscape projects part of our system. In a wider perspective we intend to publish our information system on the Internet (only for querying and not for updating or

manipulating data). A first attempt of this kind of application will be carried out during the next summer on a small excavation we started last year on the castle of Miranduolo in Chiusdino, south of Siena (the first results are published in Nardini 2001a; <http://medievalarchaeology.unisi.it/NewPages/miranduolo/mir.html>); we hope to accomplish it by the time this paper will be published.

What's still missing in intrasite GIS is the third dimension, which we started experimenting lately (or, better, we started creating the necessary supports). We do not believe in usefulness of mere 3D visualization of excavated structures; a good 3D modelling and rendering software is better than a GIS in this case.

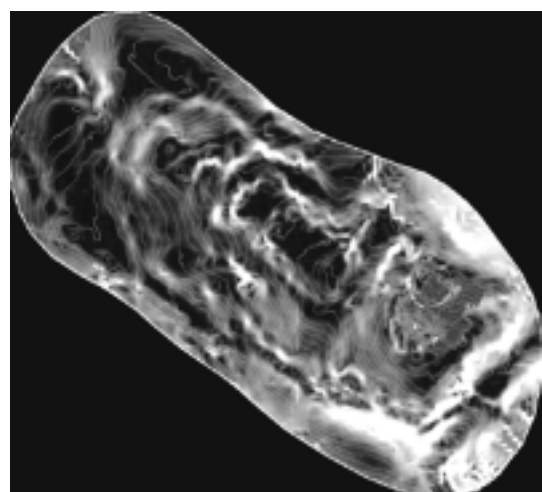


Fig. 5. Slope map applied to the hill of the Miranduolo castle of (Chiusdino, Siena).

In order to have a correct 3D assessment of excavated sites we need digital terrain models with a resolution of at least 1 meter, and an appropriate recording system during excavation (with a precision of at least 1 cm); it is our intention to experiment complete 3D recording of a small excavation area this summer. To be honest we want to make an attempt, even though it is our belief that GIS technology is not yet ready for the management of three-dimensional data (for a more in-depth discussion of this subject see the paper of Salzotti and Valenti on this same volume). To do this, we adopted a GIS software, called Geo Concept, running on Windows platforms only (our choice is due to the fact that GeoConcept works very similarly to MacMap, the GIS we use on Macintosh platforms); it supports all the basic 3D visualization techniques, implementing quite powerful spatial analysis tools, grid/dtm processing features and VRML data export.

Moreover, using this software forces us accomplish another crucial step: setting up our entire information system on Windows platform; this means avoiding any problem related to the choice of different operating systems, and especially those connected to different development stages of particular software (due to obvious market reasons it happens often, when it comes to GIS software, that updating of Macintosh versions are being released a lot of time after Windows versions).

In fact, it is our intention to completely rewrite our linking application, mainly in order to support multiple platforms (Mac OS X, based on Unix, and Windows); this should favour the packaging of a solution suitable for a large number of users within the scientific community. Even though it is premature to decide a release date, making our solution of public domain is a goal we are definitely engaged in.

References

- BOSCATO, P., FRONZA, V. and SALVADORI, F., 2000. Un archivio informatizzato per la gestione dei reperti archeozoologici. In Brogiolo, G.P. (ed.), *II Congresso Nazionale di Archeologia Medievale, Società degli Archeologi Medievisti Italiani*, Musei Civici di Santa Giulia, Brescia 28 settembre - 1 ottobre 2000, Firenze All'Insegna del Giglio.
- FRANCOVICH, R., 1990. Dalla teoria alla ricerca sul campo: il contributo dell'informatica all'archeologia medievale. *Archeologia e Calcolatori* 1: 15-27.
- FRANCOVICH, R., 1999. Archeologia medievale e informatica: dieci anni dopo. *Archeologia e Calcolatori* 10: 45-63.
- FRANCOVICH, R., NARDINI, A. and VALENTI, M., 2000. La piattaforma GIS dello scavo nella gestione di un'area cimiteriale. In Brogiolo G.P. (ed.), *II Congresso Nazionale di Archeologia Medievale, Società degli Archeologi Medievisti Italiani*, Musei Civici di Santa Giulia, Brescia 28 settembre - 1 ottobre 2000, Firenze All'Insegna del Giglio.
- FRANCOVICH, R. and VALENTI, M., 2000. La piattaforma GIS dello scavo ed il suo utilizzo: l'esperienza di Poggibonsi. In Brogiolo G.P. (ed.), *II Congresso Nazionale di Archeologia Medievale, Società degli Archeologi Medievisti Italiani*, Musei Civici di Santa Giulia, Brescia 28 settembre - 1 ottobre 2000, Firenze All'Insegna del Giglio.
- FRONZA, V. 2000. Il sistema di gestione degli archivi nello scavo di Poggio Imperiale a Poggibonsi (Insegnamento di Archeologia Medievale dell'Università di Siena). Una soluzione all'interno della "soluzione GIS". *Archeologia e Calcolatori* 11: 125-137.
- FRONZA, V., 2001. Il sistema degli archivi nella gestione di un cantiere di scavo e la sua integrazione in un sistema globale (l'esperienza senese). In Francovich, R. and Valenti, M. (eds.), *Relazioni preliminari, Workshop "Soluzioni GIS nell'informatizzazione dello scavo archeologico"*, Siena 9 giugno 2001, Siena.
- FRONZA, V., NARDINI, A., SALZOTTI, F. and VALENTI, M., 2000. A GIS solution for excavations: experience of the Siena University LIAAM. In Stancic, Z. and Veljanovski T. (eds.), *Computing Archaeology for Understanding the Past. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 28th Conference*, Ljubljana, April 2000, Oxford Archaeopress (BAR International Series 931).
- FRONZA, V. and VALENTI, M., 2000. L'utilizzo delle griglie di riferimento per lo scavo di contesti stratigrafici altomedievali: elaborazione di una soluzione informatica. In Brogiolo G.P. (ed.), *II Congresso Nazionale di Archeologia Medievale, Società degli Archeologi Medievisti Italiani*, Musei Civici di Santa Giulia, Brescia 28 settembre - 1 ottobre 2000, Firenze All'Insegna del Giglio.
- NARDINI, A., 2000. La piattaforma GIS dello scavo di Poggio Imperiale a Poggibonsi (Insegnamento di Archeologia Medievale dell'Università di Siena). Una soluzione all'interno della "soluzione GIS". *Archeologia e Calcolatori* 11: 111-123.
- NARDINI, A., 2001a. *Carta Archeologica della Provincia di Siena, Vol. IV, Chiusdino*. Siena Nuova Immagine Editrice.
- NARDINI, A. 2001b. Il modello dati nell'applicazione GIS dello scavo (l'esperienza senese). In Francovich, R. and Valenti, M. (eds.), *Relazioni preliminari, Workshop "Soluzioni GIS nell'informatizzazione dello scavo archeologico"*, Siena 9 giugno 2001, Siena.
- NARDINI, A. and SALVADORI, F., 2000. La piattaforma GIS dello scavo e i modelli distributivi di manufatti e reperti osteologici animali. In Brogiolo G.P. (ed.), *II Congresso Nazionale di Archeologia Medievale, Società degli Archeologi Medievisti Italiani*, Musei Civici di Santa Giulia, Brescia 28 settembre - 1 ottobre 2000, Firenze All'Insegna del Giglio.
- VALENTI, M. (ed.), 1996. *Poggio Imperiale a Poggibonsi (Siena). Dal villaggio di capanne al castello di pietra. I. Diagnostica archeologica e campagne di scavo 1991-1994*. Firenze All'Insegna del Giglio (Biblioteca del Dipartimento di Archeologia e Storia delle Arti. Sezione Archeologica. Università di Siena 1).
- VALENTI, M., 1998a. La gestione informatica del dato; percorsi ed evoluzioni nell'attività della cattedra di Archeologia Medievale del Dipartimento di Archeologia e Storia delle Arti-Sezione archeologica dell'Università di Siena. *Archeologia e Calcolatori* 9: 305-329.
- VALENTI, M., 1998b. Computer Science and the management of an archaeological excavation: the Poggio Imperiale Project. *Archaeological Computing Newsletter* 50 (Spring): 13-20.
- VALENTI, M. (ed.), 1999. *Carta Archeologica della Provincia di Siena, Vol. III, La Val d'Elsa (Colle di Val d'Elsa e Poggibonsi)*. Siena Nuova Immagine Editrice.
- VALENTI, M., 2000. La piattaforma GIS dello scavo nella sperimentazione dell'insegnamento di Archeologia Medievale dell'Università di Siena. Filosofia di lavoro e provocazioni, modello dei dati e "soluzione GIS". *Archeologia e Calcolatori* 11: 93-109.