

WEBGIS OPEN SOURCE SOLUTIONS FOR THE DOCUMENTATION OF ARCHAEOLOGICAL SITES

Eros AGOSTO (*), Paolo ARDISSONE (*), Fulvio RINAUDO(*)
(* Politecnico di Torino – Dipartimento Ambiente, Territorio e Geotecnologie
Corso Duca degli Abruzzi, 24 – 10129 Torino – Italy
Tel. +39-011.564.7719/7687
Email: eros.agosto@polito.it, paolo.ardissone@polito.it, fulvio.rinaudo@polito.it

KEY WORDS: GIS, WebGIS, Database, Open Source, Archaeology.

ABSTRACT:

The Land, Environment and Geo-Engineering Department of the Politecnico di Torino is involved in a research project (ARCHEOMEDSAT), with the purpose of defining the architecture and components of a suitable information system for the management of excavation data and for the provision of innovative services to the visitors of archaeological sites in the Mediterranean area. GIS solutions for the archaeological digs of Arslantepe and Elaiussa Sebaste (TK) have been conceived and developed together with the Department of Archaeological Science of the “Sapienza Università di Roma”. To guarantee a more flexible usage, both in data collection phase and in the system fruition, a supporting management system has been created based on a web architecture. The system is based on Open Source tools, in order to build a flexible and scalable architecture. Its structure is described in the paper

1. INTRODUCTION

The DITAG - Land, Environment and Geo-Engineering Department - Geomatic Research Group at the Politecnico di Torino has been involved in a research project called ARCHEOMEDSAT. The project is funded by the Italian Ministry of University and Research, and seven Italian University departments cooperate in it. The aim of the project is to define the architecture and the components of a suitable information system for the provision of innovative services to the visitors of several archaeological sites in the Mediterranean area.

This system also has to be designed to support research activities and to facilitate the management of the logistics and emergency activities in these areas.

First of all, to achieve the target of this research project, data and contents of the archaeological investigations have been organised in a data model, defining rules, conventions, file formats, able to integrate and manage them in a GIS environment.

The specific task of our research group is to conceive and set up, together with the Department of Archaeological Science of the Università di Roma “La Sapienza”, GIS solutions for the archaeological digs of Arslantepe and Elaiussa Sebaste (TK). All the data acquired during the excavations campaigns (reports, forms, surveys, photos....) has been re-organised and put in a GIS oriented data structure designed in order to analyse, preserve and share it.

2. DATA MANAGEMENT POLICY

Sometimes cultural heritage research projects based on GIS technology are exposed to the risk of facing an onerous expense which gets limited results and may fail the expectations.

The more common risks are:

- to create resources not well documented, not verifiable, not comparable with similar project.
- to constrain the GIS project to the original software platform

- to limit the GIS applications at a starting and planning level, without using it for everyday analysis

To avoid these risks the system has to be planned in order to be integrated with other GIS projects, to be loaded on different software products, to be simple and understandable not only by the authors.

We designed an approach based on using several file formats for the data exchange, on the definitions of a metadata set, and finally on following several guidelines for the project documentation.

In particular, in the designing of the GIS data model three problems had to be faced:

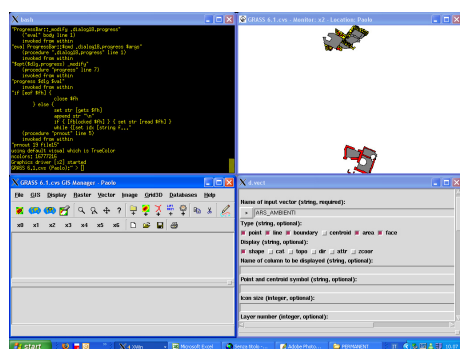
1. the different approaches, investigation methodologies and the usual data collection procedures of involved archaeological groups had to be preserved;
2. both the acquisition of new information coded according to the project requirements and the *ex post* re-organisation of datasets obtained from studies performed in the past had to be scheduled;
3. safety, conservation and portability of the stored data and of the system itself had to be guaranteed.

The different archaeological contexts and the differences in the investigation and in the usual data collection procedure suggest the question if it is possible to define a fixed GIS data structure shared by the three sites. After many discussions, the whole research group agree on the fundamental importance of safeguarding the different approaches and investigation methodologies of each archaeological group.

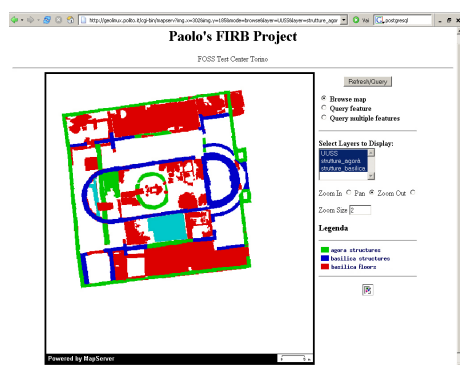
In the GIS data model designing two different tasks have to be taken in consideration: the acquisition and the codification of new information based on the project requirements, and the *ex post* reorganisation of datasets coming from studies performed in the past. Specialists and final users in fact require to process analyses of the data stored in the GIS which join or compare data obtained from different research activities. The data model was structured in order to preserve the original logic, to safeguard the correctness, and to declare the origin of the data. In general the specialists' requirements have mediated with each informative system necessity to have information organised and coded in an efficient manner. Moreover, unlike “traditional” databases, GIS technologies require the thematic and spatial information to be structured

and managed together, making the implementation task more complicated. The description and organisation of the relations and the information obtained from different investigations have to be matched by the definition of the relationships between information and spatial data.

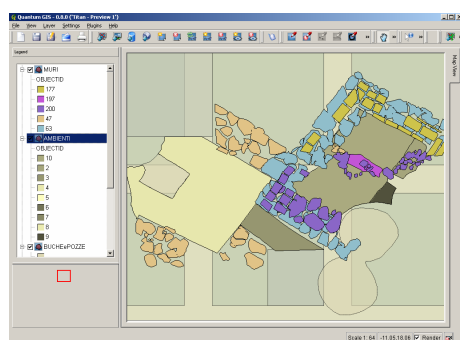
A specific GIS oriented data model adapted to different contexts requirements was planned for each of the three archaeological site, and at the same time a common technological approach was designed with the aim of ensuring longevity, preservation, sharing and spreading of the data.



Arslantepe data loaded on a GRASS Project



Elaiussa Sebaste data published using MAPSERVER



Arslantepe data loaded on QGIS

Figure 1 – Data loaded and managed on different GIS software programmes

Practically speaking, a unique data model that could be applied to each and every archaeological digs was not proposed (this solution would have created many problems from both a practical and epistemological point of view). The archaeologists were free to decide on the contents, data formats, data structure, analysis tools...in order to acquire

information suitable for the aims of their projects. At the same time some file formats, characterised by high portability, were suggested for the GIS implementation, and for the project conservation, sharing and spreading. Furthermore a metadata structure was proposed in order to describe the contents and formats of the contained data and to allow a first evaluation to be made of the resources stored in the system. In this way the designed system is able to create data that can be loaded and managed on different GIS software programmes, both desktop and Web, whether commercial or open source. Consequently other GIS projects will be able to integrate this data, and different users could understand the quality and limitations of the data.

3. TRANSLATION AND RE-ORGANIZATION OF DATA IN DIGITAL FORMAT

The documentation acquired during the excavations can be divided into two main classes: alphanumeric data and spatial data. The archaeological research groups have a long tradition in on field data acquisition (both for spatial and for the alphanumeric information), which is coded by a long used practise. It is not necessary to radically change this practise.

It was preferred to create an ER Model which could reproduce the logic of the alphanumeric documentation traditionally collected in a paper archive and define some rules, standards and guidelines for the creation of CAD files. In this way it was possible to obtain data in digital format, already modelled for the GIS project. Spatial and the alphanumeric component can be joined together in a GIS environment in a subsequent action.

3.1 Alphanumeric data

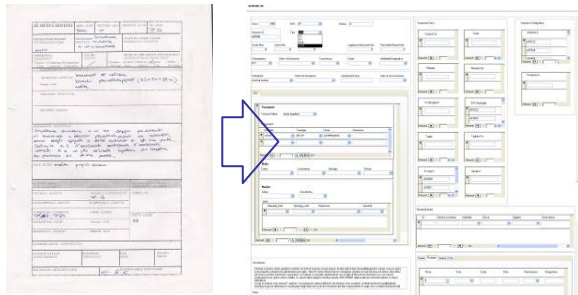
In the excavation documentation some classes of entities that could be recognised and described with a common set of attributes were pointed out. The data collection structure used by the archaeologists was followed as much as possible. Translating a data collection system based on a paper archive into a digital information system required the definition of several rules, that could be implicit and partially followed and checked. in the paper archive

A database, instead is composed of tables, made up of rows and columns, which intersect each other and create fields. The types and formats of the data that fill the fields have to be determined at the beginning of the data input: text, number, date, flags...Consequently, the person in charge of the data input has to follow fixed rules and to fill the fields with a specific type of data. This is not a drawback, it in fact guarantees the availability of data scheduled for correct and incisive analyses.

The translating from the paper archive to the database tables was an occasion to optimise several aspects of the data collection. Lists of coded attribute were used in the database structure we used as much as possible. This was for two reason: to simplify the system, in which there would be common sets of coded attributes from different tables, and to limit errors in the data input.

In the first phase of the project Microsoft Access was used to create and test the ER model and to define the coded attributes and vocabularies of the three excavations. This solution made it possible to discuss and deal with the complex phases of the translation of the data collection from the paper archive to the database. The tables and records stored in the db were also exported in a .csv file and in this

way they were ready, formalised and made available for GIS implementation.



Example of the Elaiussa "stratigraphic layer" paper form

Mask for data input

Layer	Description
1000000	Stratigraphic Layer
1000001	Stratigraphic Layer
1000002	Stratigraphic Layer
1000003	Stratigraphic Layer
1000004	Stratigraphic Layer
1000005	Stratigraphic Layer
1000006	Stratigraphic Layer
1000007	Stratigraphic Layer
1000008	Stratigraphic Layer
1000009	Stratigraphic Layer
1000010	Stratigraphic Layer
1000011	Stratigraphic Layer
1000012	Stratigraphic Layer
1000013	Stratigraphic Layer
1000014	Stratigraphic Layer
1000015	Stratigraphic Layer
1000016	Stratigraphic Layer
1000017	Stratigraphic Layer
1000018	Stratigraphic Layer
1000019	Stratigraphic Layer
1000020	Stratigraphic Layer
1000021	Stratigraphic Layer
1000022	Stratigraphic Layer
1000023	Stratigraphic Layer
1000024	Stratigraphic Layer
1000025	Stratigraphic Layer
1000026	Stratigraphic Layer
1000027	Stratigraphic Layer
1000028	Stratigraphic Layer
1000029	Stratigraphic Layer
1000030	Stratigraphic Layer
1000031	Stratigraphic Layer
1000032	Stratigraphic Layer
1000033	Stratigraphic Layer
1000034	Stratigraphic Layer
1000035	Stratigraphic Layer
1000036	Stratigraphic Layer
1000037	Stratigraphic Layer
1000038	Stratigraphic Layer
1000039	Stratigraphic Layer
1000040	Stratigraphic Layer
1000041	Stratigraphic Layer
1000042	Stratigraphic Layer
1000043	Stratigraphic Layer
1000044	Stratigraphic Layer
1000045	Stratigraphic Layer
1000046	Stratigraphic Layer
1000047	Stratigraphic Layer
1000048	Stratigraphic Layer
1000049	Stratigraphic Layer
1000050	Stratigraphic Layer

dbf file of Stratigraphic Layers data

csv file of Stratigraphic Layers data

Figure 2 - Alphanumeric data treatment

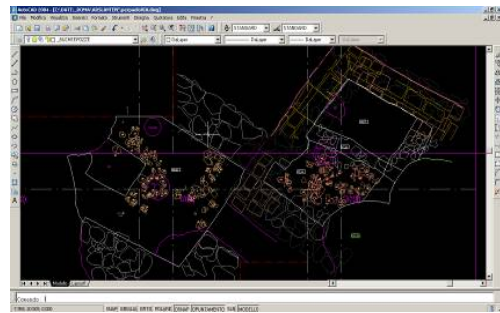
3.2 Spatial data

The graphic documentation obtained from the excavations is composed of the surveys that were carried out during the dig procedures and returned in a CAD environment. The on-field surveys data were not joined to the alphanumeric data (obviously in the registering procedure and not from a conceptual point of view).

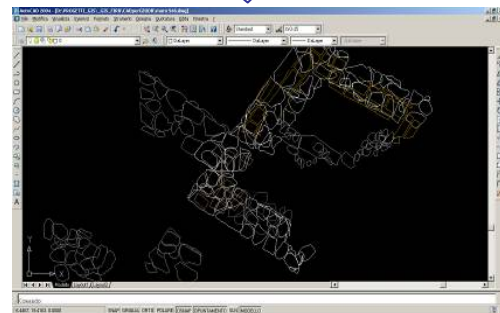
The surveys map the areas, the structures and the objects encountered during the dig, using the descriptive capability of the CAD software much more than the possibility of representing the reality using symbols, layers, or predefined forms for fixed types of objects.

The computer-aided (CAD) software programmes are conceived to help users to generate and visualise the graphic representation of the data, without adding additional information. GIS instead requires information that is structured and organised according to its meaning and its querying possibilities. Practically speaking, the CAD files created on field surveys are useful for graphic

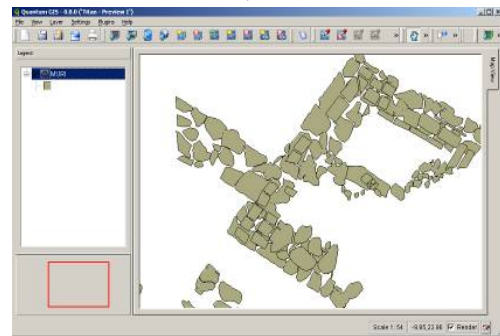
representations, but are not directly loadable in a GIS environment, and they cannot be integrated with the alphanumeric data.



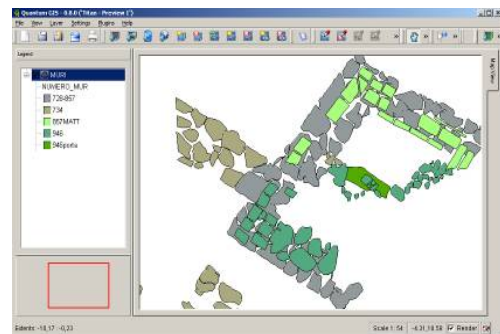
On field surveys from the Arslantepe excavation



Editing operations on the CAD file



Geometry population of the entity class using the dxf file



Entity class "Wall" in a GIS map interface discriminated by the attributes

Figure 3 - Spatial data treatment

The drawings, maps, surveys and CAD files have to be reconverted following the data model designed for the GIS. The CAD files populate the geometry of the entity classes described by the alphanumeric attributes.

CAD environments are usually not able to manage more than one attribute for each single entity. Therefore, the alphanumeric attributes are stored and managed in an external file (.csv format), creating a univocal key code to link the CAD objects to their attribute records.

The main problem is that inside a CAD file there are frequently layers composed by points, polylines and closed and open polygons. GIS software programmes are not able to understand and interpret conflicting information. They can only manage the geometry of the elements if these are described unambiguously as closed polygons, polylines or points.

For these reasons several guidelines were defined and followed in order to create CAD files that are easily imported into a GIS environment.

This problem was faced by reconsidering and converting the already acquired data, but above all by scheduling the acquisition of the new data in the future campaigns.

The conversion of the CAD data, in order to facilitate their exchange with GIS, requires long and complex editing operations. Entities have to be singled out and reorganised in specific new layers. They then have to be controlled. In fact, areas have to be represented by closed polygons. Lines have to be continuous polylines, and in general the geometry information of the entities has to be coherent with the class they are going to populate. This operation can only be done manually. At the beginning of the project we considered the possibility of programming a batch processing tool to convert the CAD file acquired on an archaeological field in a CAD file that would be suitable to easily implement in a GIS environment. However, when testing the required operations it became clear how the interpretation of the survey outputs done by the specialists is important. This necessity of a specialist interpretation made it impossible to define an automatic procedure.

4. CREATION OF A SUPPORTING MANAGEMENT SYSTEM BASED ON A WEB ARCHITECTURE.

The archaeologists required a system that could be updated daily during the excavation activities, and integrated in the post excavation phases. In order to guarantee a more flexible usage, both in the data collection phase and in the fruition of the system, a supporting management system, based on a web architecture was created.

The first and undeniable advantage of using Internet technology to publish GIS data consists in the simplicity and economy of the implementation. Users who access GIS resources, shared using a Web server and published on a specific Web site, no longer need particular hardware configurations or software knowledge: it is sufficient to connect to the WWW, and use a traditional Web browser.

The system is completely based on Open Source tools, which was considered to be the best solution to achieve the project goals. In fact, with these technologies it is possible to build complete and flexible solutions, using the continuous updating of software packets and applications and being sure about the efficiency of the tools, minimizing the starting up and maintenance costs.

The server is based on a Linux Debian operating system; Apache Server, which is the most used and powerful on the Internet network, was selected as the Web Server.

All the data was stored in a PostgreSQL DataBase Server combined with its spatial extension PostGIS: in this way, both geometric and alphanumeric information was stored in a single tool; as a consequence the system architecture is simple and light.

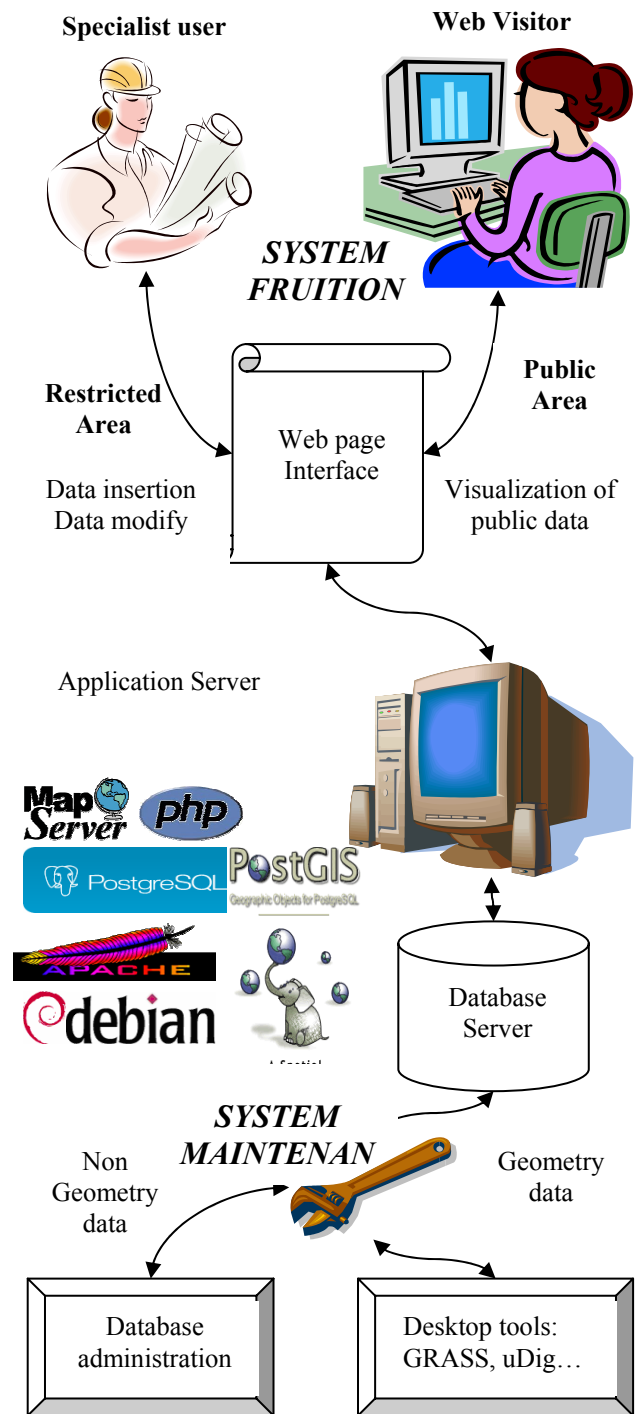


Figure 4 - System Architecture

A specific web application based on Php scripts was developed to load the alphanumeric data that were acquired during the excavation phase: a web page specific forms makes available for a correct and controlled data input. In this way the data immediately become digital and available for

web fruition; this solution also avoids problems of misalignment and duplication. Besides no plug-ins are need on the client side. Whereas alphanumeric data are uploaded via web and directly stored in the database and are therefore immediately accessible, geometric surveyed data are sent in .dxf format to the system. These files are structured following the specific directives, and spatial data are inserted manually in the database with a simple back-office work.

At present *dx2postgis*, a tool to convert DXF files to PostGIS geometry tables is being tested (<http://sourceforge.net/projects/dx2postgis/>).

Using *dx2postgis*, a single DXF file is converted to a PostgreSQL - PostGIS SQL script to create and populate tables. The connection between the features and the attributes is guaranteed by a group of structured rules for CAD data acquiring and implementation pointed out in the first phase of the project; a check procedure was however developed and a and it is implemented using a specific tool based on a QGIS (<http://qgis.org/>) interface.

then assign the univocal key code to each entity to link it to the related attribute records.

The Geometric data are published using the University of Minnesota MapServer (<http://mapserver.gis.umn.edu/>), which is an Open Source development environment for building Web mapping applications; the web user interface is based on a MapServer client open framework called Pmapper, which exploits the benefits of Ajax technologies.

The web page for the data insertion has restricted access and only authorized users can log in and operate; furthermore security settings are adopted in the firewall and Apache configuration on the server.

Each authorized user has its own account on the server and his/her operations are recorded. Suitable security settings are adopted for data storage safety: the hard disks adopt raid controller to avoid data losing in case of disk breakdown.

5. CONCLUSIONS

The application allows archaeologists to easily store, manage, analyze and share data which otherwise would be difficultly to be accessed. The system represents a framework where the results of studies and research activities carried out on the Arslantepe and Elaiussa archaeological sites can also be communicated. It should become a fundamental tool for archaeologists and other researchers, permitting complex queries and multifactorial analyses between stratigraphic and architectural data, and the information deducible from the material culture finds.

ARCHEOMEDSAT is an on going project. Now the data input operations have been approached, and we would like to use it in the every day activities of the archaeologists and researchers involved in the excavations.

The research project also shows how, after an accurate GIS data model definition, Open Source technologies can be exploited to create low-cost products, that can be customized according to the archaeologists' requirements, in order to satisfy their specific needs.

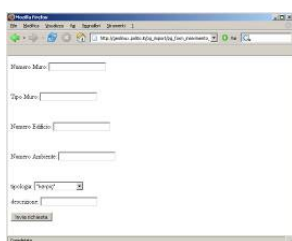
REFERENCES

AGOSTO E., ARDISSONE P., RINAUDO F., TODISCO V., 2006, Open source system for spatial and temporal data management. The botanical census of the royal racconigi park study case, Proceedings of ISPRS Commission IV International Symposium on "Geospatial Databases for Sustainable Development", Goa, India, September 27-30,2006

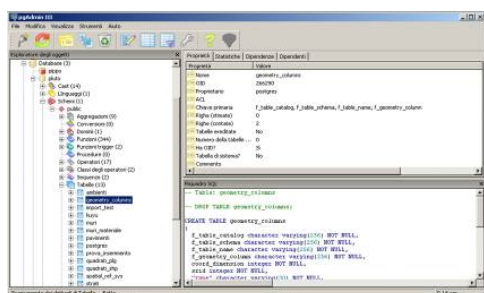
ALLEN, K.M.S., GREEN, S.W. AND ZUBROW, E.B.W. (eds) 1990. *Interpreting Space: GIS and archaeology*. London: Taylor & Francis.

ARDISSONE P., RINAUDO F., 2005, *a GIS for the management of historical and archaeological data*, Proceedings of CIPA XX International Symposium, Torino

BISWEL S., CROPPER L., EVANS J., GAFFNEY C.F., LEACH P., 1995, *GIS and excavation: a cautionary tale from Shepton Mallet, Somerset, England*, In LOCK G. and STANCIC Z.(a cura di),



Mask for data input (via web)



Tree table in a PostgreSQL client tool

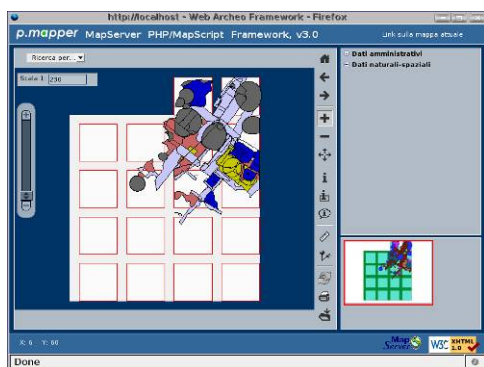


Fig 3. Supporting management system.

Using QGIS, in fact, archaeologists can connect to the PostGIS geometry table, check and edit the spatial data and

Archaeology and geographic information system: a European perspective, pp.269-285.

EQUINI SCHNEIDER E. (ed), 2003, *Elaiussa Sebaste II. Un porto fra Oriente e Occidente*, I-II, L' "Erma" di Bretschneider, Roma.

FRANGIPANE M., PALMIERI A. 1988 - *A protourban centre of the Late Uruk period, in Perspectives on Protourbanization in Eastern Anatolia: Arslantepe (Malatya). An Interim Report on 1975-1983 campaigns*, in FRANGIPANE M., PALMIERI A. (a cura di), *Origini XIII/2*: 287-454

GILLINS M., WISE A.L., 1998, *GIS guide to good practise*, Oxford.

WHEATLEY D., GILLINGS M, 2002, *Spatial technology and archaeology: archaeological applications of GIS* London.

WISE A.L., 2000, *Building theory into GIS based landscape analysis*, in LOCKYEAR K., SLY T.J.T., MIHAILESCU-BIRLIBA (a cura di), *CAA96 computer applications and quantitative methods in archaeology*, pp. 141-147.

ACKNOWLEDGEMENT

the ARCHEOMEDSAT project, coordinated by prof. Manzoni, involves seven Italian universities (Universities of Trieste, Pavia, Palermo, Milano, Cagliari, Torino, Roma). In particular authors would like to thank professors Equini Schneider and Frangipane, directors since many years of the archaeological excavations at Elaiussa Sebaste and Arslantepe. Issues described in the paper can be so broadly and completely analysed thanks to their support, guidance and collaboration.