GEOARQUEOS: A SYSTEM FOR THE CREATION, UPDATING AND VALIDATION OF THE DIGITAL CARTOGRAPHY OF THE ANDALUSIAN ARCHAEOLOGICAL HERITAGE

1. INTRODUCTION: THE MANAGEMENT OF ARCHAEOLOGICAL HERITAGE IN ANDALUSIA

Since the transfer of all cultural responsibilities from the Central Administration to the different autonomous regions of Spain in the early 1980s, the Spanish Historical and Cultural Heritage has been largely ruled by the different regional governments (QUEROL et al. 1995). In Andalusia this transfer took place in 1985, and since then, the cultural department (Dirección General de Bienes Culturales - DGBBCC) of the regional government (Junta de Andalucía) has been responsible for the protection, conservation, investigation and dissemination of the regional Andalusian Archaeological Heritage.

Thus, the IAPH (Andalusian Institute of Historical Heritage) was created in 1990 as an integrating part of the DGBBCC, in charge of research and development activities within the various domains of Cultural Heritage Management (CRM). Within the IAPH, the Documentation Centre (one of the administrative departments in which the IAPH is divided) is responsible for the creation and maintenance of systems for heritage documentation. The main task of these systems is to provide high quality information support and technical services to all social agents concerned with the cultural and historical properties of the region (LADRÓN DE GUEVARA 1996) (Fig. 1).

Since its inception, the Documentation Centre of the IAPH has mainly concentrated on the creation of the Information System for the Andalusian Heritage (SIPHA), which embraces movable objects of artistic value (MARTINEZ 1994; 1995), historic buildings (PICO 1996; MUÑOZ et al. 1997; PICO, HUMANES 1998), places and objects of ethnological interest (HERNÁNDEZ, QUINTERO 1998), the archaeological resource (GONZALEZ-CAMPOS, FERNÁNDEZ 1996; AMORES et al. 1997) as well as several forms of documentation, including bibliographic, graphic and textual.

The Andalusian autonomous region, located in southern Spain, is composed of eight provinces covering a total extension of 87,000 Km², which is over a sixth of the entire Spanish territory. Both the extension of the region as well as the wide diversity and richness of its past in both cultural and ethnographic terms, account for the complexity of the problems involved in the management of the Andalusian historical heritage in all its forms.
2. CREATING A NEED: THE INTRODUCTION OF GIS IN THE MANAGEMENT OF THE ANDALUSIAN ARCHAEOLOGICAL RESOURCE

Within the Documentation Centre of the IAPH, work on the design and implementation of an information system for the regional archaeological heritage began as recently as 1995. First, all the information available about the Andalusian archaeological sites was standardised, coded and computerised in a sites database named DatARQUEOS (GONZALEZ-CAMPOS, FERNÁNDEZ 1996; FERNÁNDEZ, MONDÉJAR, DÍAZ 1998). However, from the onset, it became obvious that a GIS platform would be an essential component of the system in order to achieve a relatively high capacity to spatially query and map all the information held in the database (GONZALEZ-CAMPOS, FERNÁNDEZ 1996).

Thus, at the same time as the data input in DatARQUEOS was being carried out, a collaboration scheme was established with the Department of Prehistory and Archaeology of the University in Seville in order to create the first digital map of the Andalusian archaeological resources. This map offered for the first time a comprehensive image of the spatial distribution of the 7555 sites which had been registered between 1986 and 1991 (AMORES et al. 1997, 1998, 1999). The source from which this map was created was not the DatARQUEOS database itself, but a previously existing set of dBase tables which had been created as part of the regional environmental information system (SinambA) by the environment authorities, but had never been utilised from the point of view of CRM.
The creation of this first digital coverage of the regional archaeological resources demanded a series of transformations in the georeferencing of the sites. This involved, for example, changing the coordinate system in reference to the grid of the geographical Institute of the Spanish Army into the standard UTM grid; similarly, all references to the UTM zone 29 were changed to zone 30, so that the entire inventory of the region was referred to a single UTM zone. Once these and other data processing problems were overcome, the resulting map (Fig. 2) enabled a preliminary overall image of the spatial distribution of the regional archaeological inventory, which offered little homogeneity.

Undoubtedly, several variables have conditioned the irregularity of the distribution, and although the grade of incidence of each variable (or set of variables) is as yet to be studied in depth, it was easy possible to point out some of the patterns underlying that specific spatial distribution.

On the one hand, there is the historiographic factor, which clearly expresses itself in the form of research preferences or trends. This is the case in several areas where archaeological surveys were made taking either cartographic (map sheets) or administrative (municipalities) boundaries as the main spatial referents. Thus, in some municipal districts there were heavy concentrations of sites whilst in the surrounding ones the concentrations were considerably less dense or simply non-existent. In some other cases it became
readily apparent that the surveyed area did actually coincide with the edges of the 1:50,000 cartography used by the archaeologists when doing the survey (Amores et al. 1998).

On the other hand there is the historical factor, that is to say, those variables affecting the distribution of the population throughout the region in each prehistoric or historic period. This involves historical and geographical factors such as the average altitude, lithology, agricultural capability, mineral resources, etc. which have affected the way human settlement has varied across the entire region in the past.

Finally, contemporary building and economic activities are increasingly influencing the particular distribution of archaeological sites depicted in the above mentioned digital map. This is the case for those patterns of land use involving heavy surface alterations which put archaeological sites at risk or simply destroy them, for example road building or massive reforestations.

The elaboration of this first digital map provided the basis for a preliminary interpretation of the territorial dimension of the region’s archaeological resources, something that had been totally impossible until now. Moreover, from the point of view of the inventory of sites to be contained in the DatARQUEOS database, the coverage also empowered further decisions about new methods and documentation policies to follow in order to gain a more homogeneous picture of the archaeological resource existing in the region.

One of the most obvious steps to be taken was precisely the creation of a second digital map, this time using the information available in DatARQUEOS, as these data offered a number of qualitative advantages, especially in the following areas:

– Firstly, it provided more accurate data collected from various recent inventorying seasons, which had been carried out since 1992. Within this set of information, archaeological sites were described and located with polygonal boundaries on a 1:10,000 scale cartographic base (González-Campos, Fernández 1996; Amores et al. 1998; Fernández, Mondéjar, Díaz 1998).

– Secondly, it held the information from former inventory records, which became fully computerised (this time by professional archaeologists using the new DatARQUEOS design) in 1997. Compared to the .dbf tables that had been used to generate the first digital coverage, the general qualification of these data was clearly advantageous because it attributed the right priorities for the correct management of the regional archaeological resource.

– Thirdly, the alphanumerical information contained in the data tables could be plotted directly on the maps by means of simple queries and searches performed in ArcView. Most importantly, each site in the map could be identified by a single numeric id which was related to the relevant information held in DatARQUEOS. Thus it became possible to graphically visualise any set of archaeological elements just by performing a SQL connection to the database.
Hence, the second phase of collaboration with the Department of Prehistory and Archaeology of the University of Seville began in 1997, its main objective being the elaboration of the second, updated digital map. This map was to include both the point-sites recorded on 1:50,000 scale map sheets (many of which had already been included in the first digital map) as well as the new polygon sites which had been located on the 1:10,000 scale maps and which contained recently updated information (Fig. 3).

This task was carried out in each province, and was not completed until July 1999 due to difficulties in the quality of the data which had not been previously foreseen. In many cases, the UTM coordinates were not as accurate as expected and had to be corrected. Furthermore, the criteria used by the different fieldwork teams to delimit archaeological sites were ill-defined and not sufficiently standardised. Only in the 1998 survey season, a set of criteria were explicitly laid so that fieldworkers used the same norms when locating sites and describing their shape. This was also the first season in which the use of computer-based records was officially established by the IAPH.

This time loss in the elaboration of the second digital map because of the georeferencing corrections proved critical for further important decisions affecting the data management policy followed within the Documentation Center. Basically, once the coverage was finished, it became apparent that the information contained in it was quantitatively and qualitatively inferior to the information already available in DatARQUEOS.
The DatARQUEOS terminology had been updated following the criteria advanced in the recently finished IAPH thesaurus of historical heritage terms (IAPH 1998). This thesaurus incorporated a wide range of new terms as well as a rationalisation of the terminology already existing (which in turn had already been associated with the new digital map) on the basis of an integrative approach to heritage management.

The range of fields contained in DatARQUEOS had been partly rearranged, especially in what concerned chronological description and functional classification of sites, in order to permit higher quality searches and cartographic output.

A significant volume of new information had already been produced by new surveying and cataloguing projects carried out by the Documentation Centre itself, the DGBBCC and the Provincial Cultural Delegations.

At this point, the need for a system capable of constantly updating the digital coverages with all the new information that is continuously introduced into the DatARQUEOS database became all the more clear. There was an urgent necessity to create a functional software application capable of bridging DatARQUEOS with the GIS maps already in production, providing a quality test of the data with sufficient speed. This application was to be inserted within the SIPHA as part of the tools developed by the Documentation Centre itself to help other governmental and private organisations in their data queries about the Historical Heritage.

3. Objectives of the Computer System

Basically, the main functional aim of the system was the periodical generation of new coverages obtained from the alphanumerical information provided by DatARQUEOS, as well as its own validation. To achieve this, GeoARQUEOS was to create a point topology coverage that corresponded to point sites, as well as a polygonal topology coverage which corresponded to polygonal sites. During this process it was necessary that the system should be capable of analysing the original data, checking for new, modified or mistaken information, then compare the newly resulting coverage with the previously existing one, to finally generate a report output. All these processes were to be carried out over a short period of time and within a user-friendly interface and working environment.

In summary, as it was designed, the GeoARQUEOS programme had three objectives:

1. Elaboration of digital coverages using the information available in DatARQUEOS.
2. Detection of georeferencing errors in the resulting coverage.
3. Automatic comparisons between the resulting and former coverages.
3.1 The elaboration of digital coverages using the information available in DatARQUEOS

The first requirement established for the system was that all the data processing should be carried out straight from the DatARQUEOS data tables without any previous alterations or changes of their information. In fact, the DatARQUEOS data table containing the site coordinates displays various coordinate formats. First of all, the coordinates refer to two different UTM zones (29 and 30). This reference to the original UTM zone in which the site is located must be kept in order to assist further fieldwork operated with map sheets; therefore, automatic conversion of all coordinates to zone 30 in DatARQUEOS was ruled out at an early stage. The application was supposed to detect the values which had been georeferenced within zone 29 and convert them into zone 30 to comply with the policy followed by other bodies producing regional digital cartography.

As was mentioned above, archaeological sites are represented in DatARQUEOS as point sets (unique pairs of coordinates) and polygonal sets (series of pairs of coordinates). This duality was to be kept in the new cartographic application in order to handle point sets and polygonal sets in the form of independent coverages, without any need for previous manual discrimination. The entities represented in the newly created coverages also contain a series of basic data extracted from the database, including the site id code, name, province and municipality, chronology, typology (functional classification), sources and legal status.

3.2 Detection of errors in the resulting coverage and comparison with previous coverages

One of the conclusions drawn from the two previous experiences in transferring the database of sites into a GIS was that the coordinates were often subject to a number of errors (some of them existing in the original sources, and some others made at the input stage). These flaws were sometimes sufficiently evident as to be easily detected and corrected. However, in some other cases they were much harder to discover. Considering the entire bulk of data involved (more than 11,000 site records), the correction process slowed down enormously the production of new coverages. The new programme, therefore, should be capable of detecting and correcting automatically those errors. They can be broken down as follows:

– Sites located outside the boundaries of their municipality. This fault is detected when the municipality associated with each archaeological site in the database does not match with the municipality in which the site is located on the newly generated coverage.
– Polygonal sites with disordered correlativity between nodes. This fault exists when those pairs of coordinates delimiting an archaeological site have been allocated at random, causing the loss of its right geometric shape.
– Polygonal sites with faulty coordinates. This is the case when a fault has been made while allocating one or more coordinates, generating unrealistic polygonal forms.

3.3 Comparison of the resulting coverage with previous coverages

The possibility of constantly updating the digital cartography of the archaeological resource was inextricably linked to the need of establishing a filtering system to compare new coverages with older ones to avoid redundancies. This involved:
– New point set and polygonal sites included on the map.
– Point set and polygonal sites eliminated from the new map.
– Point set and polygonal sites where changes related to the location and surrounding area of a new coverage have been registered.

On the other hand, the constant arrival of new archaeological data to the Documentation Centre further required the implementation of an orderly documentation record of coverages, which was to be registered and stored twice a year. This was considered necessary to keep evidence on the progression, refinement and enlargement of the regional archaeological cartography.

4. THE DESIGN OF GEOARQUEOS

The application has been developed within ArcView 3.1 using the programming language Avenue and the Visual Basic 6.0 tool. ArcView 3.1 includes a report generator, Seagate Crystal Reports, which allows reporting of errors and modifications in the new coverages generated by both the programme itself and the IPAH. The utilisation of ArcView also skips the use of Crystal Reports from Visual Basic, which would involve the installation of libraries and drivers for all users. This version also includes the Dialog Designer extension, which permits the elaboration of the system dialog boxes (Fig 4).

Since ArcView does not allow a straightforward handling of data tables, and made necessary the processing of several thousand records in each validation process, three tools have been developed using Visual Basic with Microsoft DAO (Data Access Objects) 3.5. These tools, which can be started directly from ArcView are the following:
– Genera tool. This permits the creation of two tables, Genera Puntos and Genera_poligonos directly from the DatARQUEOS table of site nodes, whilst at the same time performs a query which provides extra additional information to each site record.
GeoARQUEOS: A system for the creation, updating and validation of digital cartography

– **Vacia error tool.** This tool deletes all errors generated by the last run of the application.

– **Delfile tool.** This detects elements that have been created, erased or modified, comparing the shape files generated by the application with the ones already existing in the IAPH.

Apart from this, other tools and operations were performed to complete the system. Thus, to change the UTM coordinates from zone 29 to zone 30, a dynamic link library (dll) called *Pasohuso* has been developed in Visual C++. Also, ArcView’s normal interface has been modified and 12 new buttons or icons have been included (these allow the complete or partial use of the functions included in the system).

Other general characteristics of GeoARQUEOS are the following:

– The two tables mentioned above (*Genera_puntos* and *Genera_poligonos*) are added through a SQL connection. From these two tables, the shape files *Yacimientos_punto* and *Yacimientos_poligono* are in turn created, checking for changes in the UTM whenever necessary by means of the *Pasohuso* dynamic link.

– One of the most common errors within DatARQUEOS tables of site coordinates is the ill-defined ordering of the nodes (pairs of coordinates) of poligonally-defined archaeological sites. To detect this kind of fault, the system makes sure that none of the polygons is formed by a pair or a group of connected polygons.

– The system further detects whether a polygon may contain within its nodes any faulty digits (typically typing errors) causing its size or shape to assume absurd parameters. This is achieved through a geometric test of each polygon, since faulty polygons have an extremely low area-length relation.
– Finally, the system checks for contradictions in the georeferencing of the site and the municipality in which it is located. This is achieved by superimposing on the site coverage the entire coverage of Andalusian municipalities (boundaries) on a 1:100,000 scale.

The programme has been copied on a CD (Fig. 5), and would run on any PC which has had ArcView 3.1 installed and which contains the DatARQUEOS data tables.

5. System feeding: GeoARQUEOS and quality control

Since its creation, the archaeological sub-system of the SIPHA has been loaded with various data sources. At present, some of these sources have
been completely migrated onto DatARQUEOS whereas others are still in progress:

– Inventory of rock art sites. Completed.
– Inventory of underwater sites. Completed.
– Files on zones declared of cultural interest (high protection status). Partial.
– Files on specific inscriptions in the Andalusian heritage catalogue (high protection status). Partial.
– Reports&Projects on rescue excavations. Partial.
– Reports&Projects on systematic archaeological interventions (research programmes). Partial.
– Urban or territorial planning. Partial.
– Bilbiographic sources. Partial.
– Direct survey.

Between 1996 and 1998, DatARQUEOS was loaded at the IAPH following a schedule set by the Documentation Centre, and contained the most recent inventory of archaeological sites, as well as the inventories of rock art sites and underwater sites (the latter had been previously produced as part of a research project dealing with the anthropogenic factors of risks of coastal underwater heritage).

After this massive load, the future data input of the system is contemplated in a much more de-centralised manner because of the volume of information which still has to be normalised and entered into DatARQUEOS. Most of the information proceeds from local archaeological interventions carried out throughout Andalucia, the results of which are deposited in the provincial delegations of the Cultural Department (DGBBCC). This de-centralised input has been structured as follows:

– Projects dealt within the Documentation Centre. These are projects of documentation and research focusing on specific areas and themes and carried out under the direct auspices of the IAPH. Through them, all the available information pertaining to the archaeological resources of concrete municipalities or areas is collected.

– Archaeological surveys. The Documentation Centre lends the DatARQUEOS database to professionals who carry out archaeological surveys in order to avoid paper-based recording which would later have to be computerised. These surveys are carried out either as (1) part of more ample research projects carried out by individual bodies (i.e. universities, etc.), (2) as part of preventive measures, forming part of environmental impact evaluation of large pub-
lic works (road networks, open casts, power infrastructures, etc.) or (3) as part of urban planning revisions which include specific regulations to protect the Archaeological Heritage.

– Yearly updates of the regional inventory of the archaeological sites. These revisions are carried out by the DGBBCC using DatARQUEOS. This is the most productive and systematic input activity.

– Normalisation and input of information carried out by the provincial delegations of the DGBBCC. This work scheme is still only partially operating, due to the deficiencies in technical and human resources in these organizations. The short term objective is to provide the adequate infrastructure, so that each delegation can include in the database all the information generated in their daily routine.

Once the new information has been recorded and stored at the Documentation Centre, two filtering processes take place at the same time. First, the data are checked for errors according to the normalisation criteria established. At the same time, a digital coverage is generated with the new archaeological sites, using the GeoARQUEOS system to identify possible errors in the georeferencing and to check whether the changes introduced in the previous coverages match the established objectives. Once filtered and tested, the information is automatically downloaded into the SIPHA system, forming an updated provincial digital map. On the other hand, if faults are found in the information sent which cannot be solved by the technical staff in the centre, the information is sent back to be corrected by the team or person in charge of the fieldwork.

The results obtained from the newly created cartography (obvious faults and modifications or the inclusion of new sites on the map) are filed together with a copy of the coverages which will not be used again; in this way a historical record of the documentation process carried out will be available in the future.

At the moment of writing this article, the whole regional inventory of sites is began to revised, once the third regional coverage was created with the assistance of GeoARQUEOS. In this case, the production of a complete regional coverage plus a detailed report on obvious faults in the georeferencing has been made faster and more efficient.

6. CONCLUSION

The GeoARQUEOS programme has provided a considerable improvement in the handling of the information related to the Andalusian Archaeological Heritage. This progress has had effects principally on the process of quality control of the related information, minimising the amount of errors which escape manual filtering and get storing information in the system.
The continuous updating of the available archaeological digital cartography enhances and qualifies the Information Service of the Documentation Centre, which provides information required to users, developers, businesses, public organisations, etc. (Fig. 6), and breaks down the origin of information requests, which are mainly four: internal requests from the different centres and from professionals working in IAPH; requests from institutions such as city councils or university departments (46); personal requests from professionals as well as researchers who are not connected to any specific institution; and finally, a group of data queries from companies that deal with environmental projects or territorial infrastructures.

These data suggest that the SIPHA is increasingly approaching its explicitly stated goal of becoming a public service for Heritage Management. Such an Information Service must be capable of dynamically responding to the needs of all social agents concerning CRM, including research supports and assistance to professionals and the public in general. This follows the philosophy stated in the Andalusian Law on Historical Heritage (passed by the regional Parliament in 1991) according to which only by bringing the Cultural Heritage closer to our society, can an adequately preventive and sustainable heritage policy be achieved.

The effectiveness of the aforementioned preventive policy depends on the quality of the information available. Therefore, the Documentation Centre of the IAPH will continue to work in the computerised applications development which will help to reach this goal.

SILVIA FERNÁNDEZ CACHO  
Instituto Andaluz del Patrimonio Histórico

ENRIQUE MANUEL BLASCO ARANDA, RAFAEL NAVASCUÉS FERNÁNDEZ-VICTORIO  
Guadaltel Ingeniería de Sistemas S.A.
BIBLIOGRAPHY


IAPH 1998, Tesoro de Patrimonio Histórico Andaluz, Junta de Andalucía, Consejería de Cultura, Instituto Andaluz del Patrimonio Histórico, Peligros (Granada).


ABSTRACT

GeoARQUEOS is a programme designed with Avenue and Crystal Reports for the purpose of keeping the digital cartography of the Andalusian Archaeological Heritage updated. Its three basic functions may be summarised as follows: a) Automatic elaboration of digital coverages and data models using the information available in the Andalusian Archaeological Heritage Database (DatARQUEOS); b) Detection of geo-referencing errors in the resulting coverage (archaeological sites outside of their municipality boundaries and evident errors in the assignment of coordinates); c) Automatic comparisons between new and previous coverages, in order to offer information about the new, cancelled or modified archaeological sites in the new coverage. The GeoARQUEOS programme represents a considerable improvement in the handling of information related to the Andalusian Archaeological Heritage. This progress has been possible mainly because it acts as a means of quality control for the data by minimising the number of errors which escape manual filtering when information is stored in the system.