The 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2008)
M. Ashley, S. Hermon, A. Proenca, and K. Rodriguez-Echavarria (Editors)

Workshops, Short and Project Papers

A Systematization, Storage and Retrieval System for Archaeological Pottery

A. L. Martínez-Carrillo [1], A. Ruiz-Rodríguez [1], F. Mozas-Martínez [2], J. M. Valderrama-Zafra [2], I. Murillo del Valle [2], A. Felicetti [3]

[1]Andalusian Center of Iberian Archaeology, Jaén, Spain

[2]Department of Graphic Engineering, Design and Projects, Jaén, Spain

[3]PIN, University of Florence, Prato, Italy

Abstract

The systematization and the entry of the archaeological artefacts are essential tasks that have to carry out according with an open system capable of admitting different types of information. In this sense, is fundamental the employment of computer devices that permit the order, the storage and the search of any kind of information. This article shows an example of systematization and arrangement of archaeological pottery with the use of a database. This work is included on the CATA project (Archaeological Wheel Pottery of Andalusia in its acronyms in Spanish). The main objective of the project is the implementation of a database which is accessible by Internet, containing assorted information about pottery vessels and fragments found in Andalusia in different periods (iberian, roman and medieval).

This article describes the different steps realized in the database development: unification and standardization of the information concerning to archaeological pottery, data association in different levels of information (metric, qualitative, contextual and preservation) and the design of the tables of the database in Postgres, which is a free relational database software that allows complex searches, view of objects, and the insertion of new types of information, functions and index methods. Also is exposed the implementation of the database in Internet.

1. Introduction

Due to the large volume of ceramic material found in archaeological explorations, this material directly contributes to a great amount of our historical understanding. The study and analysis of ceramics constitutes one of the most frequent activities of the archaeologist's work. A common task is the classification of thousands of ceramic sherds which are discovered. This classification contributes to the possibility of deducing forms, functions and chronologies of the vessel fragments.

The study of archaeological pottery is a continuity of methologies used since the founding of this discipline. Although diverse procedures have been defined in the past for the classification and categorization of pottery as of yet no unified methodology exists.

This article will define a computer system that stores and classifies vessels and sherds found in archaeological interventions. The CATA (Andalusian Archaeological Pottery-Wheel Material) project is an automated implementation for classifying and categorizing the above mentioned artefacts. The system's principal objective is the creation of a referencing framework for Andalusian pottery which dates from the Iberian, Roman and Medieval periods. The three fundamental functionalities of the system are geographical referencing of artefacts, cross-section time analysis and functional categorization of artefacts.

This article is based on the following points: the creation of protocols for the study of archaeological pottery, the design and the structure of database, and the availability of the information in Internet.

2. Protocols for the study of archaeological pottery

The first step in creating the CATA methodology was the definition of the protocols required for the analysis and study of pottery vessels found in Andalusia. These protocols are valid and applicable to different historical periods.

The second step was to define the key variables which synthesis the description and definition of vessels and sherds. The four most important variables are based on the work of Orton, Tyers and Vinci [OTV93]: date, distribution, functionality and state of preservation.

CATA adds an additional variable to the above:

2.1. Measurement variables

Measurements are based on a raster or vector image that is the drawing of the pottery vessel. The following sub-classification of measurements is used:

- Basic measurements: vessel diameter, height, volume and weight.
- Complementary measurements: these measurements define and numerically specify the most significant parts of the morphology of a vessel (rim, handle and base).

The above numerical information is obtained from the measurements related to vessel image. They can be inserted manually or semi automatically. Recent developed software called Profile Analysis Tool (PAT) oriented towards the drawing of shapes, including a utility that allows semi-automatic measurements of the pottery vessel drawing [LMM*06].

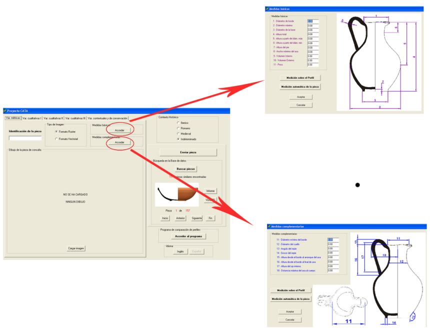

Figure 1: *Interfaces to insert measurement variables*

Qualitative variables are related to the manufacturing process of the vessel. Therefore inside this range of variables include aspects regarding the shaped of the vessel, type of oven treatment, chemical composition of the clay used and additives added to the clay. Included in the qualitative observations is the description of the morphology of the vessel distinguishing rims, handles and bases, surface treatment and decorative aspects, and chemical analysis.

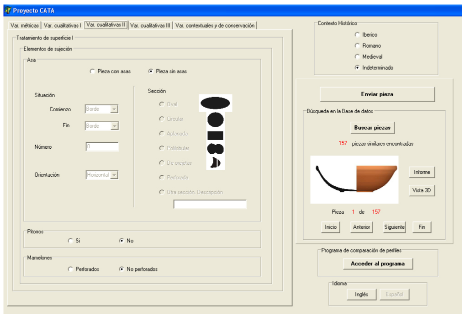

Figure 2: *Interface to insert the treatment of a vessel*

2.2. Preservation variables

The preservation variables are related to the physical state of the vessel (complete or fragmented), alterations suffered by the artefact, and the manufacturing treatments used to create the vessel (used propose to restoration and conservation of a pottery vessel).

2.3. Contextual variables

Finally, variables have been added to describe the context in which the artefact was found. A pottery vessel or fragment is associated with a temporal and spatial context. The identification of the artefact context will allow correct dating, deduction of functionality, and the application of geographical significance. In this sense the work of Shepard [SHE56] marks an inflection point in the study of archaeological pottery taking account different aspects as the chronology, the storage and distribution and the technological development. These variables identify and characterize the areas in which the ceramic fragments have been documented.

3. Design architecture of the CATA database

3.1. Index design considerations

Cultural Heritage applications are now characterized by their massive utilisation of digital media [CAC*04]. This has been employed to document sites, artefacts and restorations. Up to recently, such documentation is mostly based on pictures, reports, and analysis.

The database engine selected for the CATA system was Postgres due to the flexible indexing systems integrated into this database. The CATA system uses 2 index types: Balanced Tree (B-Tree) and General Search Tree Index (GiST).

For common queries a B-Tree is used given it's capability of producing a fairly shallow hierarchical structure for the information contained in a table. This helps to reduce the number of rows visited to retrieve a particular set of data. The Postgres B-Tree also supports the use of comparison operators and pattern matching operators (LIKE, BETWEEN,

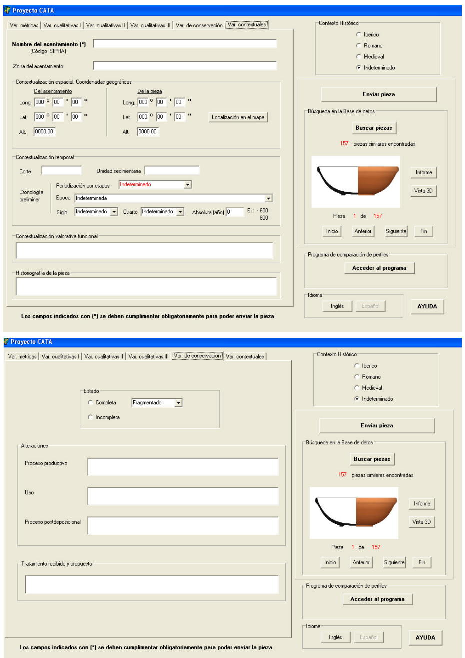

Figure 3: *Interfaces to insert contextual and preservation variables*

IN, etc.). Particular attention was paid to limit the number of columns used in composite key indexes. The largest composite index is composed 3 integer type columns which are settlement identifier, vessel identifier and global identifier. The global identifier is a wrapper used to collapse the settlement identifier and vessel identifier into one unique value. This permits a foreign key relation on only one key column in all the tables (vessel manufacturing, vessel evaluation, vessel sherds, etc...) associated with a particular vessel with the benefit of reducing indexing overhead and search times.

Given the unique nature of the CATA system which permits the search and retrieval of vessel profile geometries and associated data, the Postgres geographical information General Search Tree Index (GiST) is used to store and retrieve the vessel geometries [pos08]. Since this index requires a Spatial Reference System number (SRS) a value of -1 was introduced which means no SRS is applicable. The GiST index permits a quad based ordering of the vessel geometries according to the dimensions of the envelope required to store a vessel's profile geometry. As a result even the geometries are indexed using their vertical and horizontal measurements in order to reduce geometric query search time. The final benefit of the GiST index is associated with the true nature of its use in Postgres which is Geographical Information System (GIS) oriented index. In a future version of CATA there is anticipated the development of spatially oriented information concerning vessels and sherds.

3.2. Table schema and design

Once the above mentioned variables were defined and clarified, the next step was to create the table schemas using these variables and define the relation amongst tables.

The root of the table hierarchy is the settlement table. The first generation of master-detail tables is the following: vessel table, artificial zones and historical zones. These tables are associated with the settlement table by a unique integer relationship. The following is a description of these tables.

Vessel table: acts as a bridge between the settlements table and all information concerning sherds and vessels (quantitative and qualitative tables). It should be noted, the vessel table will be referenced against the historical zone table and the artificial zones table by its geographical relation to these tables.

Historical zones: described by usage such as housing, graveyards, battle fields, administrative areas and any human activity which may describe a zone.

Artificial zone: These types of zones are geospatial quadrants, transect, sample cuts, or any type of geometry used to organize the analysis of a settlement.

The lowest level child tables describe all the information necessary to define a vessel or its sherds (chemical analysis, historical context, etc...)

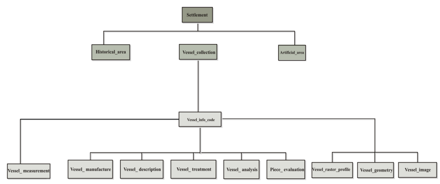

Figure 4: *Tables schema*

4. General System Architecture

4.1. User Interface

The interface is based on a single web page design which uses ActionScript 3 in order to communicate with a Simple Object Access Protocol (SOAP) server in other words a WWW2 standard design. This type of design limits the number of petitions place on the server for generating the user interface. Once uploaded to the client machine the interface acts and behaves like a desktop application. There is heavy implementation of tab layouts which permits efficient use of scarce screen real estate and is quite intuitive for the end user. The interface communicates with the server using AJAX. This implies several benefits: the interface does not lock while waiting for feed back from the server; multiple requested to the server can be sequentially generated and consumed ; only the portions of the user interface which need to be updated with new information are modified due to incoming data; and extremely low risk of attacks (virus, malware, spies, etc...) given that the client requests services using XML and receives data in XML format (only text is sent and received with no binary information). Note: there is one exception where binary information is sent and received which is the uploading and downloading of vessel profile images in PNG, JPG or TIFF formats.

4.2. Servers

Plain Old Java Objects (POJOs) are exposed by the SOAP services that compose the CATA project. The main services are the following: settlement services, artefact services and geometry services. The use of SOAP permits easy modification of existing services and scalability of new services [viv08]. The POJOs are used to interact with an Object Relational Mapping (ORM) tier which communicates with the database. The ORM tier is created using the Hibernate object persistence framework which is mapped to the CATA database [hib08]. The elimination of traditional SQL allows for queries based on object inheritance and relations amongst objects in the database (polymorphic queries) and also permits the combination of geometric queries with criteria queries (eg: give me all vessels that have similar geometry to a given vessel and that are dated between 400-500 B.C.).

5. Conclusions

In this article has been exposed the methodology followed for the achievement and storage of information regarding to archaeological ceramics.

This information has been classified in different variables (metric, qualitative, of conservation and contextual); also the indexation of geometric variables of the profile has been predicted.

The articulation of the information constitutes an open system; since this information is available in Internet.

References

[CAC*04] C C. L., AITKEN G., COTTE P., PILLAY R., PITZALIS D.: Digital capture of archive, objects and paintings for research and conservation. In *In Proc. of EVA 2004* (2004).

[hib08] HIBERNATE.ORG: http://s2kftp.CS.Berkeley.EDU:8000/, 2008.

[LMM*06] LETTNER M., MARA H., MULLER A., SABLATNIG R., SINGER M., KRENN M.: Pat: Profile analysis tool for the documentation of archaeological finds. In *In Proceedings of the 1st EVA Vienna Conference* (2006).

[OTV93] ORTON C., TYERS P., VINCE A.: *Pottery in archaeology*. Cambridge University Press, 1993.

[pos08] POSTGRES: http://www.hibernate.org/, 2008.

[SHE56] SHEPARD A. O.: *Ceramics for the archaeologist*. Washington: Carnegie Institute of Washington, 1956.

[viv08] VIVIDSOLUTIONS: http://www.vividsolutions.com/, 2008.