The Giza Plateau Mapping Project Excavation Database

T. W.W. Tonner

Giza Plateau Mapping Project, Oriental Institute of Chicago and Harvard Semitic Museum, USA

Abstract

The Giza Plateau Mapping Project which operates large-scale excavations in Cairo/Egypt is managing invaluable datasets about the building of the Pyramids of Giza. For its archives an initiative was founded to develop practical solutions for disseminating, archiving and preserving the collected information. Four key points were paramount to the development of a centralized database: low cost, flexibility, security and easy access. The internet was chosen as a platform to transport information between the members of the international research team. To keep costs down, open source software was employed featuring the operating system 'Linux', the object-relational database 'PostgreSQL', 'Apache' web server, the scripting languages 'PHP' and 'JavaScript'. Forged together, a flexible and secure web-driven database was developed integrating all archaeological datasets from the Giza Plateau Mapping Project archives.

Categories and Subject Descriptors (according to ACM CCS): H.2.4 [Systems]: Relational Database, H.2.8 [Database Application]: Scientific Databases, H.3.5 [Online Information Services]: Web-based Services, J.2 [Physical Sciences and Engineering]: Archaeology

1. Introduction

Archaeology is one of the few disciplines that must destroy its subjects to attain a deeper understanding of the processes that created them. With each excavation the original context of cultural remains is inevitably lost and replaced by an archaeological records. As the appropriate management of these records is paramount to the analysis and interpretation of the findings, information technology has become an important tool for archaeological research.

With increasing amounts of information being stored in digital collections it has become evident that not only the management but also the preservation of data has to be included into the responsibilities of archaeology.

The Giza Plateau Mapping Project (GPMP) has carried out excavations on and around the Giza Plateau in Cairo/Egypt for more 15 years. Funded by the Ancient Egypt Research Associates and under the direction of Dr. Mark Lehner (Oriental Institute of Chicago and Harvard Semitic Museum), the archaeological team consists of research specialists and a field team from numerous nations scattered around the globe that meet annually in Egypt to undertake research. Since 1999, a unique effort was made to reveal the footprint of the ‘Lost City of the Pyramids’ located about 300m south of the Sphinx. With the cultural remains dating into the reigns of the pharaohs Menkaure and Khafre, the site has provided an exceptional insight into the structure and organisation of Old Kingdom towns that provisioned the labour force for the building of the third and second pyramid on the Giza Plateau.

Large amounts of invaluable archaeological information relating to the creation of the neighbouring World Heritage monuments have been collected in the GPMP archives and the integration, accessibility and preservation of these datasets became the task of the ‘Giza Plateau Mapping Project Excavation Database’.

2. The GPMP Excavation Database

Four key points, which may apply to other large archaeological research projects, were identified to guide the development of a research database for the GPMP:

1) Cost: Database development and maintenance should be cost efficient. Archaeological projects often suffer from money shortage and cannot supply consistently high funds for ongoing software upgrades and specialist database maintenance.

2) Flexibility: The database should act as an open archive. Many research projects are running for a considerable number of years and database structure and information has to be added, edited and removed in a flexible way.

3) Security: Archaeological information is invaluable and the storage and accessibility of the collected datasets has to comply to high security standards to avoid any loss or unwanted dissemination of information.

4) Access: Easy access to the data has to be guaranteed through a powerful, but user-friendly interface. The success of a research database does rely on the intensive interaction of the digital archive with the scientists.
It became quickly apparent that only the internet would provide the easy access required. The costs of database development, archive maintenance and curation could be controlled with the application of open source software (see [JGC*04] for discussion).

The operating system dominating the open source market is ‘Linux’ which is also extensively used as the platform for providing web services through the program ‘Apache’. Additionally it hosts an extremely powerful objectrelational database system called ‘PostgreSQL’ which compares to commercial products like ‘DB2’ and ‘Oracle’. A database hosted by an internet server needs an interface that communicates between the host and the clients. This interface is provided by dynamic scripting languages such as the very popular ‘PHP’. Client interactions with a website can also be controlled with ‘JavaScript’ which complements PHP as it runs not on the server itself but on the individual users’ computer.

The combination of Linux, Apache, PostgreSQL, PHP and JavaScript was identified to be the ideal solution to develop a cost-effective, flexible, secure and easily accessible database system.

Long-term projects like GPMP require long term archiving solutions which poses several questions: Is the system flexible enough to succeed in an changing environment? Will the information be easily accessible to all involved people? How will rapid hardware and software development affect the established database and the user interface? All these concerns play into the role of digital data curation: digital datasets of archaeological information have to be maintained, revised, extended, explained and made available to the appropriate clients.

With the application of cost-effective hardware and software a balanced approach can be made to achieve the goals set out above. In case of the GPMP Excavation Database the open source environment provided a wide range of powerful application development tools which will be developed further but stay downward compatible. An ideal solution for the issue of the continuously changing hardware and software are inexpensive commercial database hosts that can be found in large numbers on the internet. As they take on the responsibility of upgrading the hardware and software that the database is running on, the risks of losing the means of accessing the data can almost be eradicated. Obviously appropriate measures for data security have to be implemented when using the internet as a base for data archiving, management and curation.

3. Conclusions

Through the development of the Giza Plateau Mapping Project Excavation Database, easy and secure access to the complete excavation archive has been achieved.

The new possibilities that this system gives to the project’s team members become clear when we consider a typical example of its daily use: while data is entered simultaneously and independently in Egypt and the US, the information stored in the database is queried and downloaded in Ireland, Poland and Japan. This unique environment where all involved people can contribute from anywhere in the world at anytime to a continuously growing archive brings the information to life.

The database and interface structure can be curated from anywhere in the world if the right access codes and software are given. Further development of the system was made easy and future-proof by only using established standard open source software packages. The costs of the development and the maintenance can be kept to a minimum which enables the project to create a long-term solution for a digital excavation data archive.

The GPMP Excavation Database has shown that the latest challenge for archaeology of archiving, managing and curating digital collections can be mastered with the right choices of software and a dedicated team. We and many other initiatives not mentioned in this article have demonstrated that easily accessible, flexible, secure and cost-effective digital archives for archaeological excavations are now a reality.

References

Design and Implementation of a Multimedia Integrated Database of Archaeological Sites on a Web Service Platform*

E. Şaykol¹, Y. Saygün¹, A. Erçil¹, A. Willis², D. Cooper², M. S. Joukowski²

¹Faculty of Engineering and Natural Sciences, Sabanci University, Orhanli, 34956, Tuzla, Istanbul, TURKEY
²Division of Engineering, Brown University, Providence, RI, USA

Abstract
Cultural heritage has been gaining more importance in the recent years in combination with sophisticated yet effective computer vision techniques. As a consequence, archaeological data, both in textual and image forms, has been considered in the development of database models. Archaeological sites are the primary source of archaeological data, and the findings are the primary targets for storage, querying and retrieval as well as exchange through appropriate mediums. This paper introduces the basic design principles of multimedia integrated database of archaeological sites on a Web service platform. Web services provide a medium for integration of various archaeological resources that are available for querying through the Web. This platform also allows archaeological data exchange via querying mechanism. The Web service platform presented in this paper can be used by both archaeologists and non-technical users to query and retrieve archaeological information through various Web accessible archaeological databases. Our work also focuses on supporting multimedia visual content-based queries for archaeological objects stored in databases. As an initial step, similarity search facility for image-based archaeological data is developed as an additional query task within our platform.

Keywords: multimedia integrated database of archaeological sites, Web service platform, effective querying, data exchange, similarity search of archaeological images.

1. Introduction
In the recent years, cultural heritage has been gaining more importance especially when considered with the perspective of computer vision and image databases. Archaeological data and archaeological sites are among the trendy applications with this respect. In order to establish a system to serve as an archaeological database, we have developed a framework (MIDAS) [SSE03] for storing vast amount of information to query archaeological data and to reconstruct the objects and archaeological sites. As our study has been continuing, instead of establishing a single database at a specific database server, we have decided to develop a decentralized platform to facilitate the querying of multiple heterogeneous archaeological databases through the Web. The main motivation behind this work lies in the fact that most of the archaeologists have their own databases for their findings on top of a corresponding data model. It is for sure that the archaeologists have various data models. Within the framework, Web services are used because they provide very useful functionalities for designing a Web accessible decentralized platform [CL04].

Due to the complex nature of the archaeological data, archaeological databases are required to store wide range of complex archaeological data. The computers mainly help the excavation process. Moreover, the data stored in the archaeological databases provides a medium for reconstruction, management, and realistic visualization. In the literature, there are studies on each of these applications as well as unified frameworks aiming at modeling the management of archaeological sites. 3D Murale [GCS*01] is an important work that models a system containing recording, reconstruction, database and visualization components. Recording tools are developed for measuring terrain, stratigraphy, buildings, building blocks, pottery, pottery sherds and statues on the archaeological site.

There exist many archaeological directories and resources on the Web, which provide browsing and searching facilities based on the content of their data. Tay Project [TAY] is an archaeological inventory on the Web. A database is created to archive the archaeological sites and the findings. The users can search the database along with the ages. Currently, the Palaeolithic/ Epipalaeolithic, Neolithic, Chalcolithic and Early Bronze Age inventories are available for database

---

*This work is supported by National Science Foundation (NSF) under the grant number IIS0205477 of the project 3D Form Models for the Representation, Manipulation, and Recovery of Shape with Applications to Archaeology and Virtual Sculpting.
searches. TheBan Mapping Project [The] in another archaeological resource site available on the Web. TheBan project has focused on Thebes, and decided to built a national database of preIslamic sites. They have built as very useful dataset and resource for Egyptologists and archaeologists in not only providing a reference and research tool, but also an aid for site management. Archaeological Resource Guide for Europe (ARGE) [Arc] contains a comprehensive set of resources for the European region and provides an extensive guide for European archaeology. The set of resources are handles as Web links, and searches can be made by text, date, subject, country, source, period, language. Compass is a database of around 5000 objects selected from the wide collection of the British Museum’s collections [Com]. The search facility provided at the site is keyword based, and the use of logical operators is allowed. Database of Irish Excavation Reports [Dat] contains summary accounts of all the excavations carried out in Ireland (North and South) from 1985 to 2000. It can be browsed or searched using multiple fields of the reports, e.g., name, title, etc. A Web service, called Mediolanum [Med], is created to facilitate international cooperation in the planning and execution of archaeological field work all over Europe. A Web site is designed for this purpose. Having selected a European region, the projects within that region can be browsed to find out a possible collaboration or information exchange. The collaboration provided at this site is project oriented.

The main contribution of this study lies in the fact that it provides a medium for archaeologists to collaborate on their individual projects as well as to facilitate information exchange on the subject matter through the Web. Besides, nontechnical users can query the system and gather archaeological information from various resources located at various sites. Another contribution is the similarity search facility for image-based archaeological data. The querying module is enriched with similarity searches based on the visual content of the image-based archaeological data, as the initial step for supporting multimedia queries.

The organization of the paper is as follows: The motivation and the design principles of Web Service Platform (WSP) are presented in Section 2 along with a simple scenario. A real-life application on Petra [JwCbCAB*98] Great Temple excavation site is summarized in Section 3. Section 4 concludes the paper and presents some future work ideas.

2. Web Service Platform

There have been studies on porting archaeological databases to the Web for an extended usage in the recent years (e.g. [JwCbCAB*98]). Along with this trend, effective querying and retrieval, handling data exchange, and managing possible collaborations are one of the primary issues for the near future. Due to the complexity of archaeological data, model variations among the archaeological resources, we have designed a Web Service Platform (WSP) for Web-accessible archaeological databases to allow effective querying and retrieval enriched with similarity search facilities.

Web services are the services offered through the Web by the help of Webbased protocols [CL04]. Web services enable the participants to communicate with each other even if they are using different information systems, platforms, etc. A Web service is registered at a predetermined URL. It accepts requests from an application using the SOAP (Simple Object Access Protocol) protocol on top of the HTTP. The requests are processed in two ways depending of the type of the request:

1. If the request is a remote procedure call to activate a method on the server side, then the return value of the serverside method is sent to the application as the response.
2. The request can be an XML (Extensible Markup Language) document based on a schema agreed prior to the request. Then, it is directly processed, which is called document oriented processing.

A Web service can make itself available to potential clients by defining a Web Services Description Language (WSDL) document (a kind of signature). A WSDL description is an XML document that gives all the pertinent information about a Web service, including its name, the operations that can be called on it, the parameters for those operations, and the location of where to send requests (endpoint). XML is a markup language that makes data portable, by proving a standard way of data exchange. A Web client can use the WSDL document to discover what the Web service offers and how to access it. In short, a Web service is a server application that implements the procedures which are available for clients to call (e.g., a database query).

A simple illustration of WSP is shown in Figure 1. Querying Server accesses the Repositories of the Archaeologists (i.e., archaeological databases), via asking to the Registry Server for UDDI (Universal Description, Discovery, and Integration project) check. Each archaeologist’s local computer runs a Web service that is able to retrieve information from its local database by executing query procedures.

The Querying Server, via connecting to these registered Web services, retrieve information from the local archaeological databases. In order to submit a query to the

![Figure 1: Web Service Platform (WSP).](image-url)
databases, it employs the Java API for XML based Remote Procedure Call (JAXRPC) [ABC*03], which is used for developing and using Web services. The Querying Server uses *stubs* for remote procedure calls. Stubs are classes that represent a service endpoint on the client. This allows a JAXRPC client to invoke a remote method on a service endpoint as though the method were local. To employ stubs, the signatures of the procedures available for a remote call have to be known in advance (and they are available through the WSDL documents of the Web services). An RPC based Web service is a collection of procedures that can be called by a remote client over the Internet [ABC*03]. The Web interface of the Querying Server is designed by Java Servlets, which are known as server side programming units for Web based access. Separate servlets have been developed based on the querying facilities that the registered Web services provided.

### 2.1. A sample scenario

A simple scenario to query within Web Service Platform (WSP) can be summarized as follows: Possibly by a Web browser, assume that a Web user initiates a querying facility at the site of the Querying Server (QS). The servlet at QS responsible for the initiated querying facility triggers a set of operations. First, QS communicates with the Registry Server (RS) to get the locations of the registered archaeological databases within WSP. Then, QS executes the database querying procedures at registered archaeological sites remotely. This communication is based on SOAP (Simple Object Access Protocol), hence the communication is paused between the two parties until the response of a request is generated. Each Web service responds to QS based on the query they received. The next operation at QS is to combine these partial results and to present them to the Web user. For the sake of simplicity, not all of the attributes of the archaeological objects are displayed at the client side. However, if the Web user wants to explore the details of a query result (e.g., a pottery found at Sagalassos excavation site), a specific servlet at QS requests the detailed information from the corresponding archaeological database, and presents it to the Web user.

### 3. A real-life application: PETRA Web Service

Petra is a famous archaeological site in Jordan, and Great Temple excavations by Brown University have been directed by Prof. Martha Sharp Joukowsky [JwCbCAB*98]. A brief information on the Petra Great Temple Excavations can be found at http://www.brown.edu/Departments/Anthropology/Petra/. The archaeological database has the following catalogs: *Archaeological Fragments, Coin, Grosso Modo, Cat, Glass*, and *Image*. All of the catalogs except Grosso Modo have one level of information. Grosso Modo catalog contains two levels; Grosso Modo Items as the first level, and related materials for a selected item as the second level.

The Web service that we have developed for Petra Great Temple excavations have Web based graphical user interfaces for querying each of the catalogs. A prototype of the system can be accessed at http://hendrix.lems.brown.edu:8080/qserver/index.html. Each querying interface have various attribute selection parts, which can be used either separately or in combination while querying the corresponding catalogue. Figure 3 shows the query specification window for coin catalogue within WSP.

Figure 2 shows the first screen when WSP is first initiated by an HTTP connection. As seen from the figure, Petra is listed as the registered archaeological database. The querying facilities that Petra provides for the Web users are listed below in the figure. Each of the six catalogues can be queried by using the specific GUIs separately (e.g., Figure 3 for Coin catalogue). Additionally, we have provided multiple query interfaces with respect to some common attributes in all of the catalogues (e.g., *year, trench*).
Figure 4 presents the graphical user interface for the specification of multiple queries among Petra Great Temple catalogs. As shown in the figure, year and trench attributes can be combined by AND and OR logical operators for more focused queries. Year values can be selected from a combobox, however trench values have to be entered by the Web user because of the data range and the representation of the trench values. The Web user selects the catalogs that he/she wants to query, and QS sends the multiple query to the selected catalogs only.

3.1. Evaluating WSP by multiple web services logs

In order to evaluate the performance of the Web Service Platform (WSP) with multiple Web services, we have installed each catalogue of Petra Great Temple to a different machine. On top of each machine, we have deployed separate Web service designs, hence set up an environment of 6 Web services. In this design, the multiple query execution facility becomes more meaningful, hence gives better ideas on the performance of WSP when multiple Web services are registered to the system.

Within this environment, the six catalogs listed in Figure 2 can be queried separately by creating separate JAXRPC connections between QS. For a single Web service search, QS sends the query formulated by the specific GUI for the catalog directly to the selected catalog. However, for multiple query by year, trench, or their combination, QS sends the query to all of the catalogs (6 Web services that are registered). QS presents the partial results in a comprehensive manner to the Web user. Figure 5 shows an excerpt from the output of the multiple query of ‘selecting objects from all of the catalogs excavated in 1996 and at trench 16’. The query returned satisfying objects from each catalog, and they are presented to the Web user.

3.2. Similarity searches in coin catalogue

Coin Catalogue of Petra Great Temple is selected as a sample dataset for similarity queries because of the fact that the items in this dataset are generally associated with coin images whose visual content can be queried within WSP.

The QS, which lies in the heart of WSP, communicates with the registered Web services to respond to the userspecified queries. The definition of how to process the data stored in a registered database is specified at the Web service side, where the database is also located. A query submitted to the system via accessing the graphical user interface at the Querying Server is transmitted to the Web services, and the results are gathered based on the definitions at the Web service side. Although it is possible to exchange messages containing image data between the querying server and a Web service, a local database is designed and implemented at QS for similarity queries. This local database acts like a caching mechanism since it stores a copy of the image data of the registered databases. A query indicating a similarity search is processed directly at the querying server, which will also reduce the transmission time significantly. This design is also more reliable for this set of queries because of being local than other types of queries requiring access to Web services.

A simple yet effective module is implemented to respond to similarity queries based on the visual content of Coin images. The coins associated with an image data are inserted in a combo box dynamically by connecting to the database at the time of the initiation of the browser window shown in Figure 6.

Having selected a coin catalogue number from the combo box and pressed the ‘Similarity Query’ button, QS starts processing the query. The visual content of the coin images is pre-processed at the time of the population of the local database at the querying server. Color and shape feature vectors are extracted from the visual content of the coin images [SGU04], and the query coin image is exhaustively searched with the rest of the coin images. The relevance ordered list of coin images is returned as a result of the query.

The color vector is a probabilistically-weighted variation of color histograms. The shape vector is a combination of two vectors: the first one is based on the angular distribution of the pixels around the centroid of the object. The second vector is the accumulation of the pixels in the concentric
circles centered at the centroid of the object [SGU04]. This approach has been used for the content-based retrieval of historical Ottoman archives successfully in [ŞSG04].

A sample similarity query is presented in Figure 7. The color and shape content of the query coin image (96-C-48) are compared with the color and shape content of the other coin images, and 4 coin images are retrieved. In the module, the similarity values have to pass a threshold, which is set as 0.80, to be listed in the result presentation window. The coin numbers in this final window serve as pointers to actual coin values stored in databases at Web services. By clicking on a coin no link, the system starts communicating with the corresponding Web service to present the details of the actual coin record.

4. Conclusion and future plans

In this paper, we have presented a Web Service Platform for Archaeological Databases having Web-access. Along with the trend in porting archaeological databases to the Web for an extended usage, effective querying and retrieval, handling data exchange, and managing possible collaborations are identified as one of the primary issues for the near future. Due to the complexity of archaeological data, data model variances among the archaeological resources, we have designed a Web Service Platform (WSP) for Web-
accessible archaeological databases to provide a medium for effective querying and retrieval. The querying module is also enhanced with similarity searches based on the visual content of the image-based archaeological data.

We are planning to extend our platform by including image, video clips, and dense-data 3D laser scans in the database. We are also planning to provide additional geometric search facilities, especially based on 3D shape model of the archaeological artifacts. Another future direction is to widen WSP by more registered archaeological databases. Our evaluations show that the performance of the system is promising when there are multiple Web services registered. We are also planning to introduce ‘semantic’ queries to WSP based on the data model and semantic structure of the registered databases (e.g., searching for complete objects).

Acknowledgement

We are grateful to Professor Benjamin Kimia for his useful suggestions throughout the project.

References


[Arc] ARCHAEOLOGICAL RESOURCE GUIDE FOR EUROPE: http://odur.let.rug.nl/arge_bin/w3-msql/entry.html.


[Com] COMPASS: http://www.thebritishmuseum.ac.uk/compass/.


