A Web Service Platform for Web-Accessible Archaeological Databases*

Ediz Şaykol[1], Yücel Saygın[1], Aytül Erçil[1],
Andrew Willis[2], David Cooper[2], Martha S. Joukowski[2]

[1]Faculty of Engineering and Natural Sciences, Sabancı University,
Orhanlı, 34956, Tuzla, Istanbul, TURKEY
[2]Division 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 a Web service platform design for various archaeological resources that are available for querying through the Web. This platform also allows archaeological data exchange. 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 visual content-based queries for archaeological objects stored in databases. As an initial step, similarity search facility for image-based data is developed as an additional query task within our platform.

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. To establish a system to serve as an archaeological database, we have developed a framework (MIDAS) [1] for storing vast amount of information to query archaeological data and to reconstruct the objects and archaeological sites. 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

* This work was supported by National Science Foundation (NSF) under the grant number **IIS-0205477** *3D Form Models for the Representation, Manipulation, and Recovery of Shape with Applications to Archaeology and Virtual Sculpting.*

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 platform [2].

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 [3] 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 [4] 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 searches. TheBan Mapping Project [5] in another archaeological resource site available on the Web. TheBan project has focused on Thebes, and decided to built a national database of pre-Islamic 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) [6] contains a comprehensive set of resources for the European region and provides an extensive guide for European archaeology. The set of resources is handled 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 [7]. The search facility provided at the site is keyword based, and the use of logical operators is allowed. Database of Irish Excavation Reports [8] 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 [9], is created to facilitate international cooperation in the planning and execution of archaeological field work all over Europe. Having selected a

European region, the projects within that region can be browsed to find out a possible collaboration or information exchange.

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, non-technical 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 [10] 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. [10]). 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, data model variances 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.

A Web service can make itself available to potential clients by defining a Web Services Description Language (WSDL) document (a kind of signature) [2]. 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),

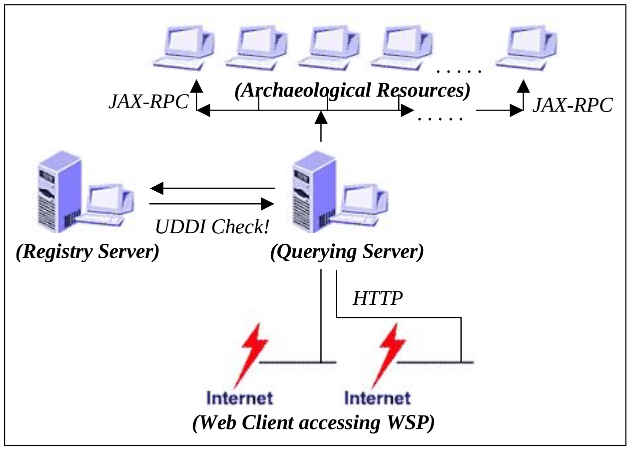


Fig. 1. Web Service Platform (WSP).

via asking to the Registry Server for UDDI 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. To submit a query to the databases, it employs (JAX-RPC) [11], 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 JAX-RPC 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 [11].

2.1 A Sample Scenario

A simple scenario to query within Web Service Platform (WSP) can be summarized as follows: 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, 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 [10]. 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.

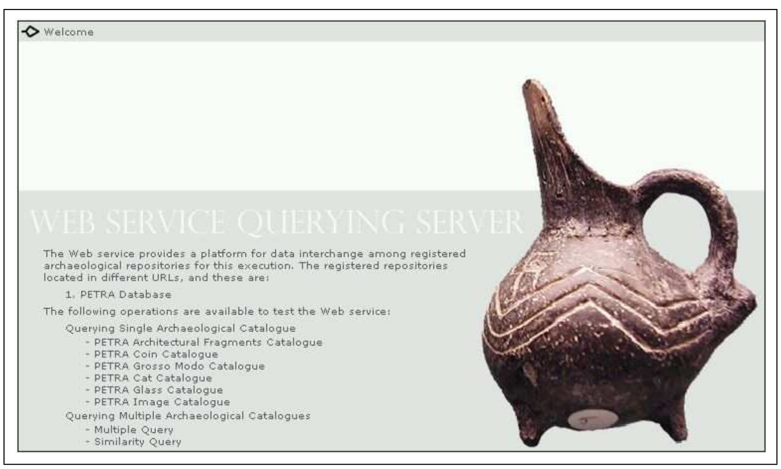


Fig. 2. The Welcome Screen of the Querying Server.

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*).

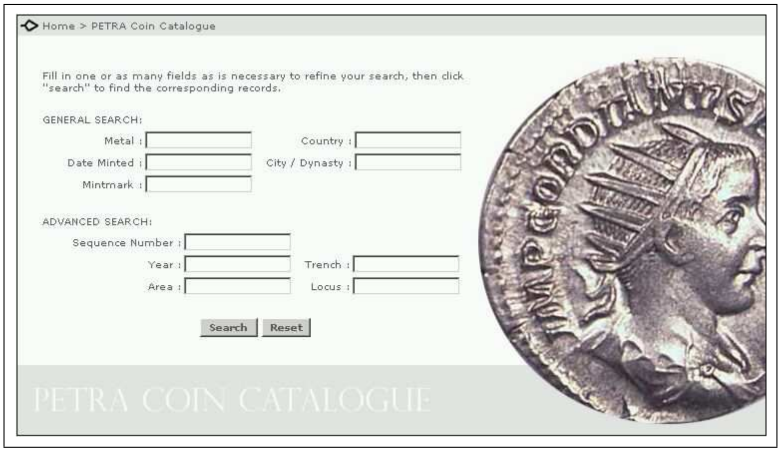


Fig. 3. The GUI of Coin Catalogue for Querying within WSP.

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 Brown University Web site `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 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 combo-box, 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.

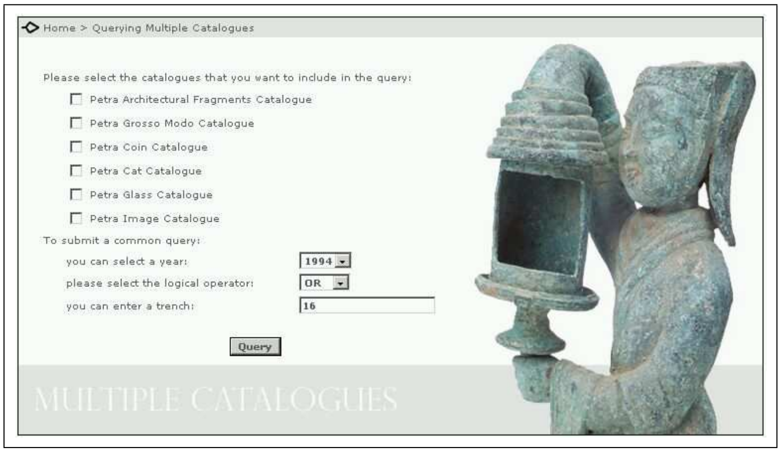


Fig. 4. Multiple Query Specification GUI among the Catalogs.

3.1 Evaluating WSP by Multiple Web Services

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 JAX-RPC 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, which is list of catalogue id's (e.g., coin no for coin catalogue, fragment no for architectural fragments catalogue) grouped by catalogue names.

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 user-specified 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 5.

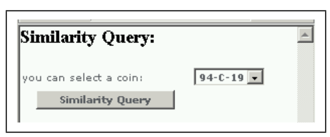
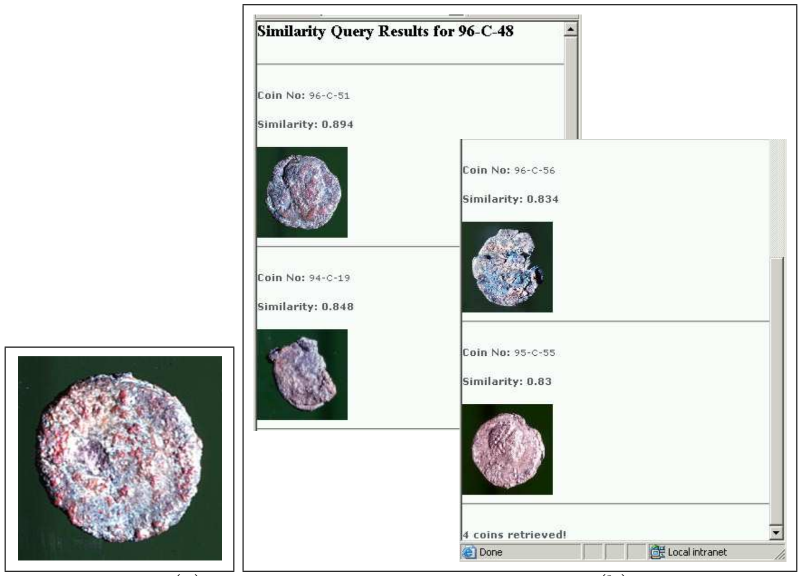


Fig. 5. The Similarity Search Window for Coin Catalogue. The coin image to be queried is selected from the combo-box to initiate the querying process.

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 [12], 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



(a) (b)

Fig. 6. The Output of Similarity Search for Coin Image 96-C-48. (a) The coin image 96-C-48. (b) The output having 4 coins retrieved, which is shown in two browser pages to the Web user.

object. The second vector is the accumulation of the pixels in the concentric circles centered at the centroid of the object [12]. This approach has been used for the content-based retrieval of historical Ottoman archives successfully in [13].

A sample similarity query is presented in Figure 6. 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 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).

References

1. Saykol, E., Saygın, Y., Erçil, A.: MIDAS: A multimedia database for archaeological sites. In Arnold, D., Chalmers, A., Niccolucci, F., eds.: 4th Int. Symp. on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST 2003), Work In Progress. (2003)
2. Castro-Lean, E.: The web within the web. IEEE Spectrum **41** (2004) 42–46
3. Grabczewski et al., E.: 3D MURALE: Multimedia database system architecture. In: Int. Symp. on Virtual Reality, Archaeology and Cultural Heritage, http://www.brunel.ac.uk/project/murale/ (2001)
4. TAY Project: (http://tayproject.org/)
5. TheBan Mapping Project: (http://www.thebanmappingproject.com/)
6. Archaeological Resource Guide for Europe: (http://odur.let.rug.nl/arge_bin/w3-msql/entry.html)
7. Compass: (http://www.thebritishmuseum.ac.uk/compass/)
8. Database of Irish Excavation Reports: (http://www.excavations.ie/pages/homepage.php)
9. Mediolanum: (http://archeonet.cilea.it/archeosite)
10. Joukowsky et al., M.: Petra Great Temple Volume 1: Brown University Excavations 1993-1997. Petra Exploration Fund, Providence, RI, USA (1998)
11. Armstrong et al., E.: The Java Web Services Tutorial. (2003)
12. Saykol, E., Güdükbay, U., Ulusoy, Ö.: Integrated querying of images by color, shape, and texture content of salient objects. In Yakhno, T., ed.: Advances in Information Systems (ADVIS'04), LNCS 3261. (2004) 363–371
13. Saykol, E., Sinop, A., Güdükbay, U., Ulusoy, Ö., Çetin, A.: Content-based retrieval of historical Ottoman documents stored as textual images. IEEE Transcations on Image Processing **13** (2004) 314–325