Research Article **Open Access**

Integration of Heterogeneous Biological Data Sources via Developing Web Service Platform

Mohd Kamir Yusof*

Fakulti Informatik, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

Abstract

Biological data is important for medical institutions or centers. The biological data used in medical institutions or centers for specified purposes such as patient report or documentation, disease statistical, etc. Medical institutions or centers must have a good platform to manage their data sources that contain biological data. In this research, Web Service Platform (WSP) was implemented for accessing heterogeneous biological data. Heterogeneous biological data means data sources that contain biological data and allocated at different places. In WSP, all data sources were integrated. The purpose of this integration is to help medical institutions or centers access all biological data. In WSP, XML is used for searching and retrieving process. XML can improve web query processing and display the information to web user in short time. JAVA technology was used in WSP development. The application was tested. The result indicates WSP can used and implemented in medical institutions or centers for accessing heterogeneous biological data.

Keywords: Biological data; Heterogeneous biological data; Web service platform

Introduction

Medical institutions or centers used information system to run their daily operation. Most of medical institutions or centers have their own web based application. Why web based application is needed in medical institutions or centers? Web based application can used anytime and anywhere as long as internet connection is available. The web based application can help these medical institutions or centers manage their works effectively and efficiently. The web based application stored lot of biological data in database server. Biological data is important in medical area for specified purpose such as patient documentation, disease presentation, statistical documentation, etc. Web based application is divided into two components. First is web application server, and second is database server. The web application server contains application sources code and database server contains biological data. Web application server is connected to database server in order to search and retrieve biological data based on queries from web users. The biological data is lot and phenomenal rate. The issue in biological data is heterogeneous biological data sources. How to integrate? Why this integration is needed? This integration is needed to ensure all web based applications can access all biological data sources. The purpose of this integration heterogeneous biological data sources is to help medical institutions or centers get more information about patient records, disease statistical, etc. A Web Service Platform (WSP) is proposed to solve this problem. The main function in WSP is to allow web based applications access all data sources. For instance, web based application A can access data sources from web based application B. The WSP will act as server between platform A and platform B. In this research also, XML approach will implement in WSP in order to improve web query processing in searching and retrieving process from heterogeneous biological data sources.

Related Work

A Web Service Platform has been applied in archaeological domain [1]. A framework (MIDAS) [2] was developed for storing vast amount of information. The function this framework is to query archaeological data and to reconstruct the objects and archaeological sites. A decentralized platform also was developed. Main function the decentralized platform is to facilitate the querying of multiple heterogeneous archaeological databases through Web. Usually, archaeological databases are required to store wide range of complex archaeological data. A computer is mainly helping the excavation process. However, data in archeological databases provide a medium for reconstruction, management, and realistic visualization. 3D Murale [3] is used for recording, reconstructing, database, and visualization components. The recording tools were developed for measuring, terrain, stratigraphy, buildings, building blocks, pottery, shreds, and statues on the archaeological site. One of archaeological projects is Tay [4]. This project was developing for archaeological inventory on the Web. A database is created to store and archive the archaeological sites and the findings. This project also provides function is for browsing and searching archaeological data. Web users can search and get information about archaeological easily through this system. Figure 1 shows interface for Tay Project.

Another projects in archaeological is The Ban Mapping [5]. In this project, national database of pre-Islamic sites was developed. They have built very useful dataset and resource for Egyptologist and Archaeologist. Web users can search information by entered data such as text, date, subject, country, source, period, and language. The web users also can enter keyword or logical operators for searching purpose. About 5000 objects were selected for the wide collection of the British Museum's collection and stored in database is called "Compass" [6]. Database of Irish Excavation Reports [7] contains summary accounts of all the excavations carried out in Ireland from 1985 to 2000. Through this database, searching and browsing process can used multiple fields of the reports such as name, title, etc. The most popular

web service is called "Mediolanum" [8]. This web service is created facilitate international cooperation in the planning and execution of archaeological field work all over Europe. Through this literature, a Web Service Platform has a potential to implement in other domain. In this paper, the WSP is implemented in biomedical. The WSP is needed in biomedical area. Through this platform, web users can search and get biomedical data especially biomedical image.

Web Service Platform (WSP)

This section presents a Web Service Platform (WSP) in accessing heterogeneous biological data.

Suppose M = {xi, x2, x3, xn}, where x represent registered web based applications and stored in register server, A. M is a member of A, this is denoted M ∈ A. Equivalently, this formula also can write A ⊇ M, read as A is a superset of M. Here, WSP is important. The importance of WSP is to ensure web user got the right or related information after searching and retrieving processes. In this research also, WSP was designed to support different programming language in web based application development. Figure 3 shows how WSP work and support different programming language in web based applications development.

In Figure 3 represents three web based applications used different languages in development. Computer A use a Java language, computer B use a PHP language and computer C use a ASP language in web based application development. WSP become as a client by defining a Web Description Language (WSDL) document [9]. Based on figure 3, first step 1 in WSP is receive query from web users either computer A, computer B or computer C. After that this will translate into a XML document. The most important component in XML is WSDL. A WSDL is description is an XML document that gives the information about a Web service such as its name, operations, and parameters for those operations, and the location of where to send requests. XML is a markup language that makes data portable, by proving a standard way of data-exchange. After translation a query to XML was done, WSP will communicate with biological data sources to search and retrieve
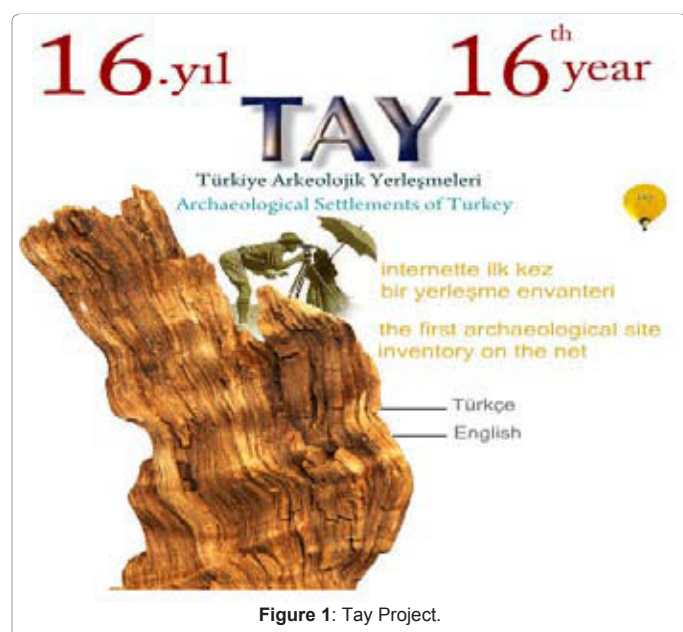

Server A represents as a register server. The register server keeps all web based applications address. Server B represents as a querying server. This server will serve a query from web user through Web Service Platform (WSP). The processes involve once web user request information as follow:-
i. Web user sends a query
ii. Server B receive a query
iii. Server B checks website address from server A.
iv. Server A will reply list of registered web based application (e.g. www.medical.com) related with biological data.
v. After server B get feedback from server A, server B will search and retrieve data from biological data sources.
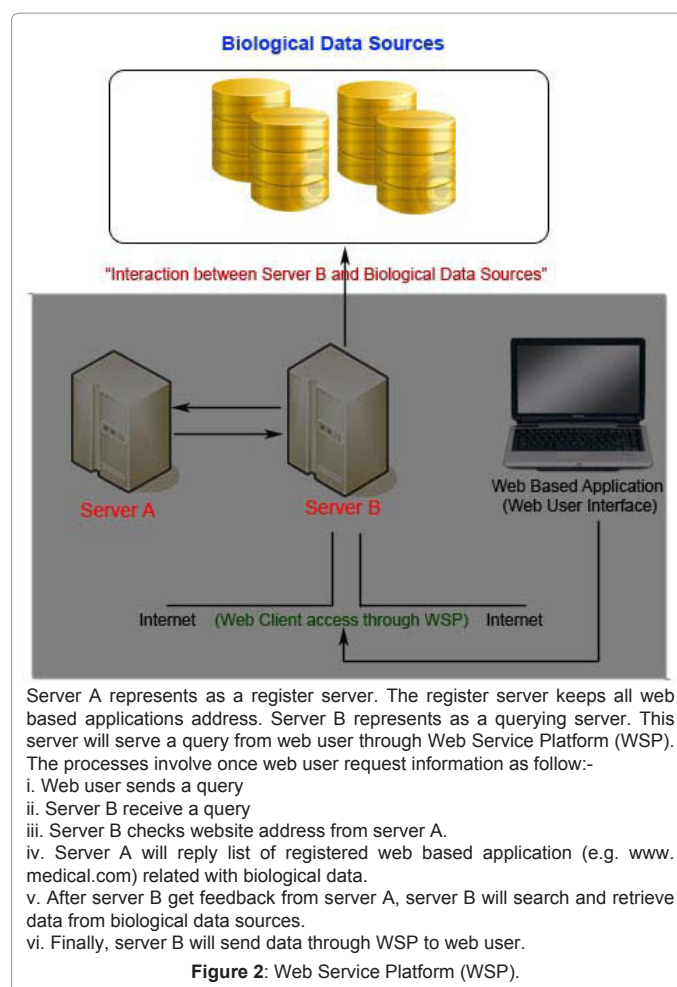vi. Finally, server B will send data through WSP to web user.

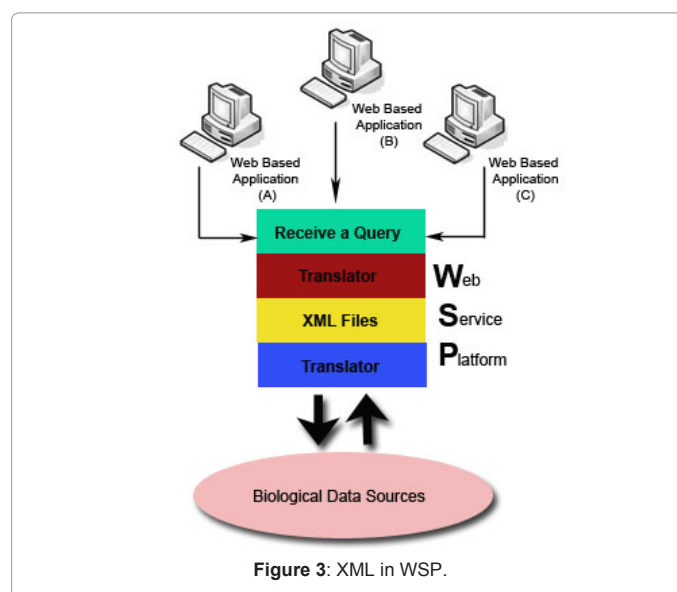**Figure 2**: Web Service Platform (WSP).

Figure 3: XML in WSP.

the data. Suppose a computer A sends a query about "*bone*". WSP will receive a query form computer A. Then, the WSP will search and retrieve data from biological data sources. After that, the WSP will send a result to computer A.

Figure 1: Tay Project.

Experimental Result

Application for WSP was develop using JAVA technology. This technology was choosing because of high performance in application development. This application was tested in order to analyze their performance for accessing heterogeneous biological data.

Experiment 1

Different web based applications with same platform (use same programming language in development) to access biological data from heterogeneous biological data sources.

Figure 4 shows interface for web service querying server. Web user will send a query to Query Server through WSP.

Figure 5 shows interface for searching. Web user is needed to enter information to text field, etc. After that, user is needed to press button "submit". Query server will receive a query from web user. Then, the query server will communicate with query server. The purpose of this communication is to know list of registered biological data sources before searching and retrieving process. After that, query server will search and retrieve data from the registered biological data sources. Finally, query server will send the results to web user through WSP.

Figure 6 shows interface after searching and retrieving process. The results will show to web user.

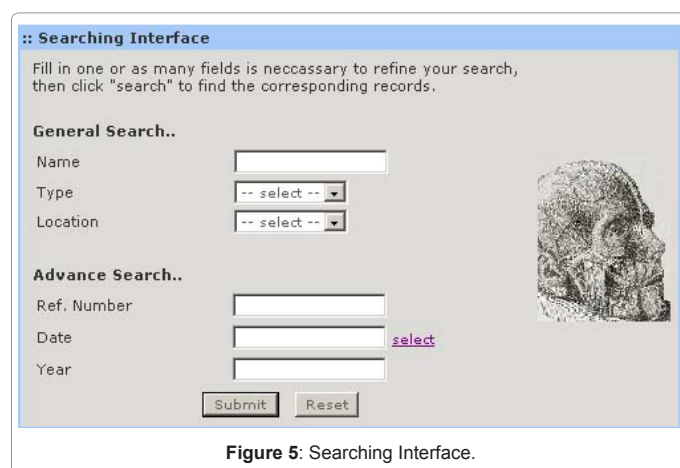**Figure 4**: Interface for Web Service Querying Server.

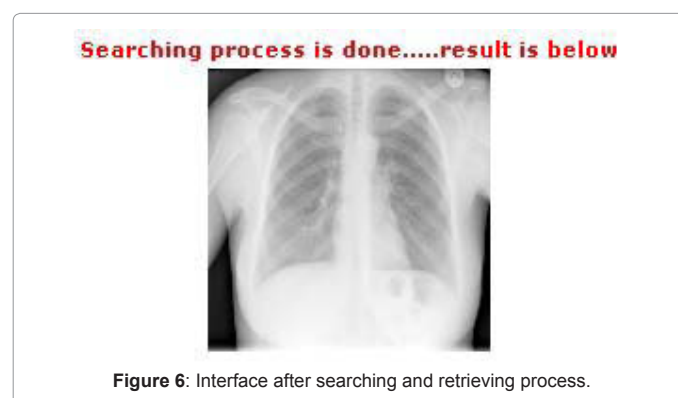**Figure 5**: Searching Interface.

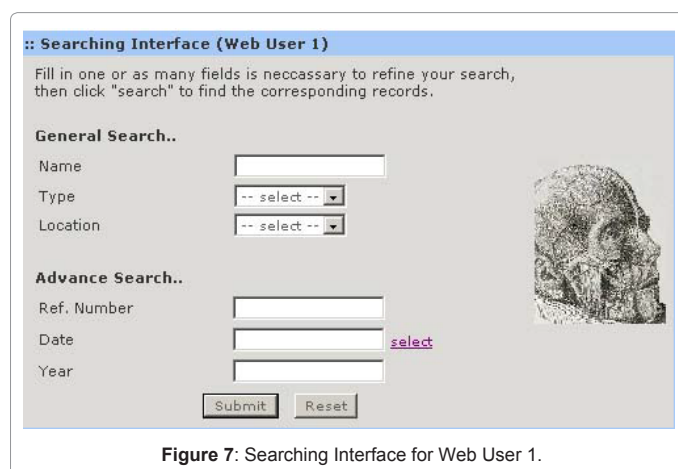**Figure 6**: Interface after searching and retrieving process.

Figure 7: Searching Interface for Web User 1.
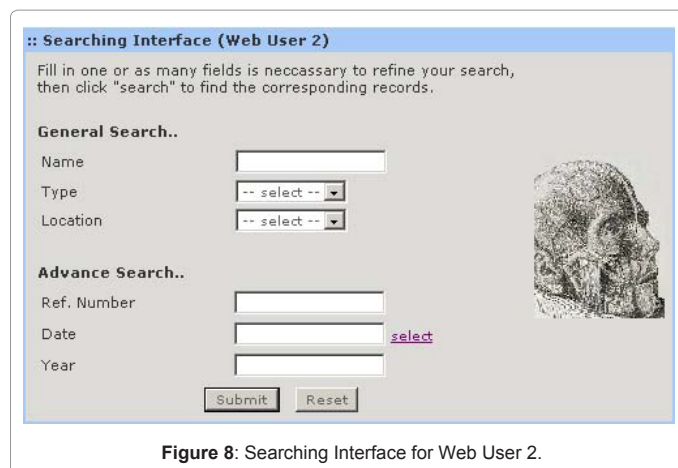

Figure 8: Searching Interface for Web User 2.

Experiment 2

Different web based applications with different platform (use different programming language in development) to access biological data from heterogeneous biological data sources.

Figure 7 shows interface for web user 1 which uses java programming language.

Figure 8 shows interface for web user 2 which uses PHP programming language.

Figure 9: Searching Interface for Web User 3.

Figure 9 shows interface for web user 3 which uses asp programming language.

Suppose all three web users entered reference number equal to "M3204". Query server will receive a query from three web users through WSP. Steps for search and retrieve are same with experiment 1. The query server will communicate with register server. After get feedback from register server, the query server will search and retrieve data from biological data sources. Finally, the three web users with different programming language will get the result through WSP.

Based on experiment 1 and experiment 2, WSP can apply for accessing heterogeneous biological data sources. WSP also can support different platform web based application (use different programming language in development) to access heterogeneous biological data sources.

Conclusions

In this research, a Web Service Platform (WSP) for heterogeneous biological data sources was developed. The capability of WSP was tested by several experiments. In these experiments, biological data sources from different location were integrated and WSP was used as a platform for web based applications to access these data sources. XML is part of component in WSP was implemented in this research. XML is good approach for data exchange, data storage, etc. In this research, XML was used to search and retrieve data from heterogeneous biological data sources. Based on experiments was done in this research, XML was able to improve web query processing during searching and retrieving process especially involved with lot of biological data. In conclusion, WSP can help medical institutions, medical centers, education institutions, research institutions, etc to manage their biological data effectively and efficiently. Biological data is important for these institutions or centers especially to provide reports such as patient statistics, patient documentation, disease documentation, etc. Hopefully, this research can help these medical institutions or medical centers more proactive in managing their information system for biological data.

Acknowledgement

References

1. Saykol E, Saygin Y, Ercil A, Willis A, Cooper D, et al. (2005) "A Web Service Platform for Web-Accessible Archaelogical Databases". ISCIS, LNCS 3733: 362-370.

2. Saykol ES, Ercil Y (2003) "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.

3. Grabczewski E, Murale D (2001) Multimedia Database System Architecture. In: Int. Symp. On Virtual Reality, Archaeology and Cultural Heritage.

4. TAY Project.

5. TheBan Mapping Project.

6. Compass.

7. Database of Irish Excavation Report.

8. Mediolanum.

9. Castro-Lean E (2004) The web within the web. IEEE Spectrum 41: 42-46.

10. Hsu W, Antani S, Long LR, Neve L, Thoma GR (2009) A Web Based Image Retrieval System for Large Biomedical Databases. Int J Med Inform 78S: S13-S24.