March 2009

Databases: From Paper-based to Web-based

Alireza Isfandyari Moghaddam
Islamic Azad University, Hamedan Branch, ali.isfandyari@gmail.com

Library Philosophy and Practice 2009

ISSN 1522-0222

Databases: From Paper-based to Web-based

Alireza Isfandyari Moghaddam
Department of Library and Information Sciences
Islamic Azad University, Hamedan Branch
Iran

Introduction

In the history of information, humans have moved relatively rapidly from an oral storytelling tradition, to manuscripts, to printed books, to filing systems, to databases. Databases were created to solve problems with file-oriented systems (Sol, 1998). The rapid growth of networked information resources and information representation have led to a reassessment of tools and techniques for information management. World Wide Web technologies play a central role in redesigning information management tools (Shiri and Revie, 2000).

Organizing and retrieving information in a databases with a limited volume of information is fairly straightforward, but as the Internet emerged, information storage and retrieval changed radically (Isfandyari, 2005). As a result, a new generation of databases called WBDBs have been created to meet user needs. Although some research such as Doldi, et al. (2005) shows that the web has reached a level of maturity in regard to scientific and qualitative content and can be considered a worthwhile source of scientific information, the present article emphasizes the important role web-based databases (WBDBs), free and fee-based, play in better and relevant information retrieval.

What is a database?

A database is a collection of data that is organized for easy storage and access These include paper-based tools like dictionaries and libraries of print materials. Computerized databases have existed for decades, and online databases are a product of the earliest days of the Internet. A web-based database (WBDB) is an organized listing of web pages (Nicholson, 2002). Online databases and WBDBs are widely available to library patrons in the entire world, and many patrons can tap into these databases from their own computers (Falk, 2005). According to Doe (2004), WBDBs are collections of information that we use all the time. In another viewpoint, Garman (1999; quoted in Xie, 2004) names online WBDB systems, such as Dialog, Lexis-Nexis, the original, or ultimate search engines, and believes that the search engines of today owe much to these originals.

History of Databases

Databases have their origin in efforts at office automation by IBM and other companies in the 1960s and 70s, resulting in models and technology that are still in use (CERN, 2000). Databases grew from an early database management system (DBMS) in the early 1960s, through networked and hierarchical relationships for data, to SQL-based and relational database models that are in use today (Zaki 2002)

Client/Server databases

Most databases in use nowadays are relational databases, but client/server databases are the basis for WBDBs. They are set up to operate 24 hours a day and are used by ISPs as well as individuals (Sol, 1998). The most common language used by relational databases is SQL.

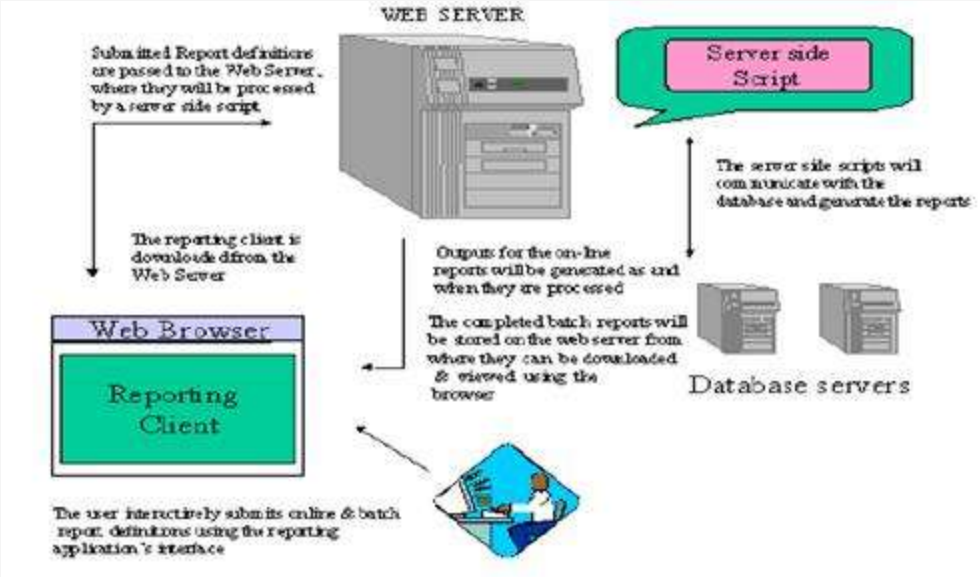

Figure 1: Activity cycle of a WBDB (Adopted from Ravindran, 1999)

Functionality of WBDBs

Feiler (1999; quoted in Wyllys, 2003) distinguishes four purposes for WBDBs:

- Web publishing
- Web data sharing
- E-commerce
- Database-driven websites

Search features included in WBDBs

Most WBDBs offer an advanced or "power" search that lets users be specific (Tarleton State University Libraries, 2004). The following images show examples of this type of search interface:

Frequently, databases also provide other ways for users to limit (narrow) searches. A few standard search limiters (restrictors) are listed and described below:

- Scholarly/Peer-Reviewed -- limits results to items from academic journals
- Document Type -- limits results to a specific type of item (for example, abstract, article, book review, editorial, report, and so on)
- Full Text -- limits results to items that are available online in full text format.
- Publication Date -- limits results to items in a specified time period.
- Publication Type -- limits results to items from specific types of publications (for example, books, newspapers, periodicals, primary sources, and so on).
- Language -- limits results to a specific language

Components of developing and maintaining WBDBs

WBDBs must be developed and maintained. The methods involved can be highly technical. The following is a summary of those methods.

- Underlying all WBDBs is a relational database-management system (RDBMS), together with one or more relational databases (RDBs) that actually contain the data or information of interest.
- A web page defined in HTML or Dynamic HTML (DHTML) controls the visual display that the user of the WBDB sees and
- An interface receives information from the user and passes it to the RDBMS, extracts information from the RDB (with the assistance of the RDBMS), and provides the information to the web page, whose HTML or DHTML structure makes the information visible.

Although Microsoft Access 97 (and later versions) offers built-in support for WBDBs, the limitations of Access restrict it to websites that experience low use (no more than a few simultaneous users). Large and heavily-used WBDBs typically use high-level RDBMSs such as IBM DB2, Informix, Microsoft SQL Server, Oracle, and Sybase. A substantial majority of such sites use Oracle.

The interfaces used for WBDBs fall into two broad classes. The first is interfaces intended for a specific application and written in a scripting language that conforms to Common Gateway Interface (GCI) standards. A script is a set of programming-language statements, typically (but not necessarily) short and used to accomplish certain actions on the Internet. Languages used in scripts include C, C++, Java, Perl, and Visual Basic for Applications. The second is interfaces developed commercially for a certain class of applications. Commercial interfaces include those of Oracle, SAP, and Siebel.

ColdFusion is a tool for aiding in the development of CGI scripts. ColdFusion's role in working with CGI scripts as analogous to that of Dreamweaver and Microsoft FrontPage in aiding the preparation of HTML-formatted pages. ColdFusion can do things such as:

- Insert and update records in database tables with HTML forms
- Submit database queries that can then be used to dynamically generate Web pages
- Intermix the results of queries with HTML tags and text for complete control over how data is displayed and formatted
- Track users and customize their view of Web pages by using information about their browser, location, or other preferences
- [Use] advanced data input and reporting features
- Validate form field entries as integer, floating point, date or numeric range
- Make conditional statements (if . . . else branching) to dynamically customize output returned to users and decisions about queries submitted to the database
- Embed SQL statements in templates to specify queries. SQL statements may be dynamically customized using data from submissions, URL query strings, and CGI environment variables, as well as the results returned from other queries
- Execute multiple SQL queries and send SQL queries to multiple databases for each client request

- Support Java and JavaScripts
- Support web browsers cookies for state control (Wyllys, 2003).

Conclusion and recommendation

In spite of the advantages of existing databases, WBDBs have become the standard in new database applications. The reasons for their popularity include:

- ***Ease of use***: point and click simplicity
- ***Accessibility***: securely accessible from anywhere
- ***Lower total cost of ownership***: to secure, scale, deploy and maintain (LightSpoke, 2003).

It seems safe to predict that WBDBs will become increasingly prevalent and increasingly sophisticated in the ways they receive information from users, as well as in the ways in which they display information to users. Though the most sophisticated uses of WBDBs are currently found in the commercial arena, such uses will spread steadily to the non-profit and other less well-financed arenas. Obviously, the Internet has added whole new dimensions to databases, not only to the type and breadth of information available- but to the complexity and problems involved in getting the information. The Internet can be seen as gigantic database or databases within a database (Doe, 2004). It is role of library practitioners to identify Internet-based services, especially WBDBs, and provide users with information on how to use them. In addition, further research is needed to assess usability, functionality, and effectiveness of WBDBs as information storage and retrieval tools.S

References

CERN (2000). A brief history of databases. Available at:http://wwwdb.web.cern.ch/wwwdb/aboutdbs/history/industry.html .

Doe, C. G. (2004). A look at web-based databases and search tools. *MultiMedia & Internet Schools 11*(5)

Doldi, L. M., et al. (2005). The web as a free source for scientific information: A comparison with fee-based databases.*Online Information Review 29*(4): 400-411.

Falk, H. (2005). State library databases on the Internet. *The Electronic Library 23*(4): 492-498.

Hundie, K. (2002). A guide to selected and freely searchable bibliographic and full-text databases on the web. *Collection Building 21*(1): 13-21.

Isfandyari, M. A. (2005). *A comparative study of overlapping and ranking of search results in metasearch engines and their common underlying search engine*s. (LIS MA thesis, Ferdowsi University of Mashhad, Iran).

LightSpoke (2003). Web-based database service: Benefits of web-based database service for business web applications. Available :http://www.lightspoke.com/web-based-database.jsp

Nicholson, S. (2002). Introduction to web databases. Available: http://www.askscott.com/sec1.html .

Shiri, A., & Revie, C. (2000). Thesauri on the web: Current developments and trends. *Online Information Review 22*(4): 273-279.

Sol, S. (1998). Introduction to databases for the web. Available: http://www.databasejournal.com/sqletc/article.php/1428721

Tarleton State University Libraries (2004). Search features offered in databases. Available: http://www.tarleton.edu/~library/library_module/unit7/7search_lm.html.

Wyllys, R. E. (2003). LIS 386.13, Information technologies and the information professions. Available: http://www.gslis.utexas.edu/~l38613dw/powerpoint/38613WebDatabases.ppt .

Xie, H. (2004). Online IR system evaluation: Online databases versus web search engines.*Online Information Review 28*(3): 211-219.

Zaki, M. (2002). CSCI-4380 database systems lecture 1. Available at: www.cs.rpi.edu/~zaki/cs4380/lectures/lecture1.ppt

Acknowledgements

Special thanks to Dr. F. Babalhavaeji for her helpful guidance.