

A Blueprint for Applications in Enterprise Information Portals

Chyan Yang, Kuang-Ming Chen, Jun-Wei Yu, Chen-Hua Fu,^{*} Szu-Hui Wu, Fan-Pyn Liu

Institute of Information Management

National Chiao Tung University

Hsinchu, Taiwan

chyan_yang@yahoo.com, fu@rs590.ndmc.edu.tw^{*}

Abstract

Electronic Commerce (E-Commerce, EC) is thoroughly changing business models of organizations (governments, corporations, and communities) and individuals the way of living and working. However, the major success will accrue to those companies that are willing to transform their organizations and business processes, which is the scope of e-Business. An Enterprise Information Portal (EIP) provides real time information and integrated applications to knowledge workers, employees, customers, business partners and the general public as well. Effective applications of EIP facilitate high quality strategic decisions. That is, an EIP can enhance an organization's productivity, improve the collaboration to facilitate E-Commerce and gain competitive advantages. However, the EIP solutions are usually too expensive to small businesses. With Enterprise Application Integration (EAI) approach, this paper presents an economic way to design a low-cost EIP that leverages existing systems. Moreover, a prototype is implemented to show the feasibility.

For the external data access, the web mining technology is utilized to mine some relevant and valuable web contents from the Internet and put these contents into the document warehouse. By combining the textual information inside the document warehouse and the numeric data from the data warehouse, competitive advantages can be provided over those who work with just the numbers.

Keywords: Enterprise Application Integration (EAI), Enterprise Information Portal (EIP), E-Commerce, Workflow, XML, Web Mining, Document Warehouse

1. Introduction

Intel Corporation president and CEO Craig Barrett said that if the Taiwanese industries want to speedup Internet growth and E-Commerce cooperation that are the first an imperative duty. He thinks that the ASPs don't exist after 5 years because all companies become ASPs. Figure 1 shows a prediction of E-enterprise's growth by IDC. Taiwanese businesses could survey outsourcing of IT and businesses refocus on overarching business objectives such as customer satisfaction, the core of competition ability and increased competitive advantages, as it controls only the

technologies critical to those programs. July, 2002. According to the IDC report, the EIP solution software not only increases, but also tendency of obvious growth. It estimates that the investment from five hundred million to three thousand million between 2001 and 2006. The growth of Internet technologies has unleashed a wave of innovations that change the way business is conducted. These shows demand immediate attention for E-Commerce.

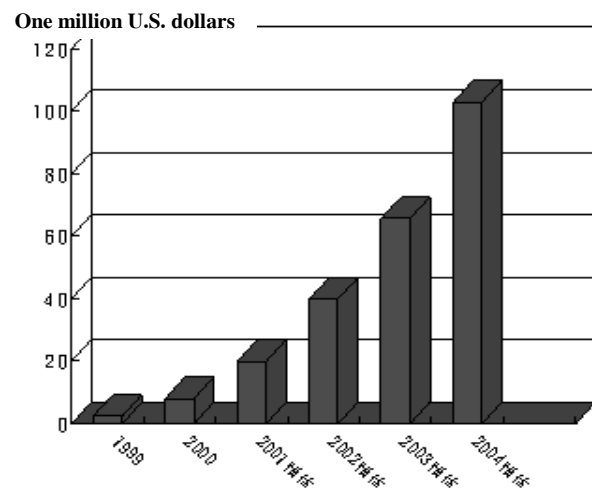


Figure 1 A prediction of e-enterprise's growth by IDC
source: IDC

A corporate portal, also called an EIP, is generally defined as a personalized, single point of access through a Web browser to critical business information located inside and outside an organization. Some refer to a "Web top," rather than browser, to include handheld devices, mobile phones, PDA, ad-hoc network devices and other Internet appliances [21]. Portals gather information from one or more servers, as well as from the Internet, and deliver that information through a single, consistent interface. This gives users one interface to access all the required documents, e-mail, Web sites, competitive information, databases, and so forth for their jobs. It's also important to view the portal not only as an internal tool in an enterprise, but as one used by the extended enterprise, including partners, suppliers, investors, and customers. The portal is a basis of any e-business strategy. The EIP

^{*} The Graduate School of National Defense Information, National Defense management College, National Defense University, Chunggho, Taipei Taiwan

solution provides a single point of access to all structured and unstructured enterprise data. It is a concept for a Web site that serves as a single gateway with a company's information and knowledge base to employees, customers, business partners, and the public as well. The EIP effectively connects users to content in context, enabling true enterprise agility [1, 17, 20].

There are two key concepts to define an EIP: access and integration. Moreover, there are two problems of information management: "Isolate of Information" and "Island of automation". Thus EIP plays a critical deployment role in reengineering legacy systems. No matter systems reengineering or business reengineering can play complementary roles when performed in an integrated. They have been stereotypically and perhaps unfairly described as existing software systems whose plans and documentation are either poor or nonexistent. Generally, legacy systems are large and/or complex; developed using unstructured data with little programmatic support. A business has to handle several kinds of issues in building brand new systems or integrating some legacy system into an EIP system [12]:

- an efficient communication and project management among teams and organizations;
- Encouraging "contextual collaboration" or ad-hoc team communication around specific data sources;
- Defining, illustrating and managing the complex interactions among the players;
- Capturing and archiving the knowledge and interactions of all players in an enterprise such that it can be used as essential data in a knowledge management/decision support mode

In the backend, data mining technologies provide analysis data tools for knowledge workers, employees, customers and business partners.

The goal of this paper is to integrate all of the legacy system into a web-based interface, the prototype is provided to preview related technology, identify open issues, and demonstrate how to convert existing client/server systems into the Web. It is therefore organized as follow. In Section 2, some of the related features of the system description are described. A brief introduction of our prototype is in Section 3. The prototype implementation for Taiwanese small business is presented in this section. The contribution of this research, technical challenges, and future work are summarized in Section 4. Finally, conclusions are made in Section 5.

2. System description

The best way to embark on an EIP system is to select the target users and understand what information they need. The information can fall into a number of categories, including:

- unstructured content (documents, hypertext content, etc.)
- structured content (data stored in relational or other types of databases)

- news (either free or for-fee)

- groupware application data (such as Lotus Domino or Microsoft Exchange)

With EIP, information can be easily organized, searched and packaged, users can friendly select what they want to access.

Our EIP system provides recommendation engines. Even if you aren't familiar with this term, you might have already experienced the functionality on E-Commerce sites like Amazon. At Amazon, after a selection you made, a recommendation list of best sale books, CDs, or movies are provided for you. This recommendation function is a useful feature to stimulate users' purchases [13].

2.1 Enterprise Information Portal (EIP)

Comparing with the MyYahoo!, a primary function of an EIP is to aggregate content from disparate sources [14]. As most companies have an environment of disparate legacy systems, application, processes, and data source, which typically interact by a maze of interconnections that are poorly documented and expensive to maintain. An EIP has quickly become important business infrastructure as they aggregate disparate data sources and attempt to unify and organize information for decision support. The integration of EIP and collaboration technology allow the organizational communication processes to be unified into single information source creating a complete and usable archive of all transaction, including the data, discussion about the data, and the decision-making process.

In addition to creating a knowledge archive, the chaotic and ad-hoc nature of B2B communication demands definable, demonstrable, and repeatable processes which companies can use to conduct business models.

2.2 Web Mining and Business Intelligent

How does user deal with gargantuan information repositories for every bit of business trivia? Moreover, the sources of data were of the old realm of business: point-of-sale terminals, inventory databases, transaction records. Attempting to understand the data, several analysis tools are used: statistical tools, OLAP systems, and Data Mining. Data Mining is the best of these tools. It is natural for one would want to analyze this data with the best data mining techniques available. The results of the data mining - the rules which say which customers are likely to buy what products at the same time, or who is about to switch to your competitor [4, 8, 9,11].

When customers visiting a web site, they provide their information based on the content of the web site. There is so much information is interesting, such as: which links customers click, what kind of web site contents they stop by longer, which terms they use to search, and when they browse. Some customers might fill out a lifestyle survey or provide names and addresses information to the visiting web site. Complex content also contains important information, such as words in articles, job descriptions and resumes, and features of competitive or complementary

products. Usually, all information is stored in a database. This results in a lot of information available on Web sites, but web users probably can not use it in a best way. To solve this problem, web-mining techniques can be used to find buried patterns in databases, and report or act on those findings.

3. System architecture

Our primary goal is to implement a web-based portal, called MyPortal, with a web-mining technique. Basically MyPortal would integrate business's legacy systems, personalized interface, and existing heterogeneous database [19]. Figure 1 shows the system architecture of MyPortal.

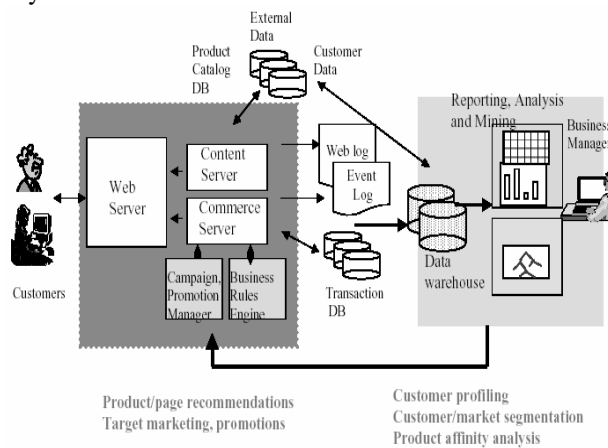


Figure 2 MyPortal system architecture

Figure 3 depicts a mapping of N-tier architecture and its implementations in MyPortal. N-tier architecture would simplify the system development environment and maintain the system easily.

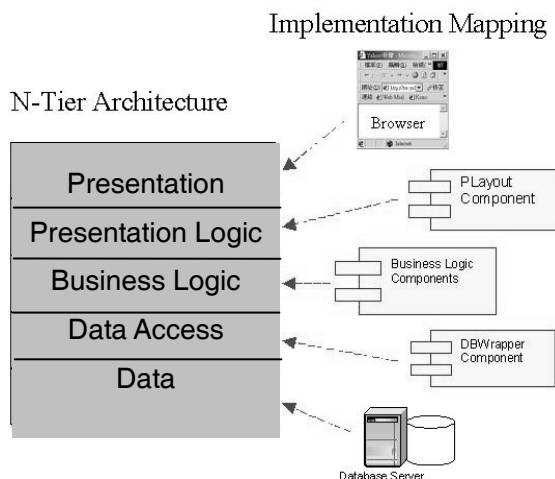


Figure 3 MyPortal mapping of N-tier architecture and its implementation

3.1 System Environment

The implementation platform is MS-window-based personal computer. In MyPortal, programming tools cost around one hundred thousand NT dollars that is affordable to most Taiwanese small businesses. System configuration

is listed as the following:

- System Environment:
 - Operation System: MS Windows 2000 Server or later
 - Web Server: MS IIS 5.0 or Apache HTTP Server
 - MS SQL Server
- Application Environment:
 - Integrated Development Environment: MC C++, ASP
 - MS FrontPage 2000
 - Java Server Page
 - JAVA 2

3.2 System Implementation

Figure 4 shows the prototype of MyPortal application. There are two major parts of the whole MyPortal: The portal and WMIS (Web-Mining Information System). MyPortal provides a personalized web-based environment and supports a web-mining technique. The WMIS provides general contents and news contents [6].



Figure 4 shows the prototype of MyPortal application

3.3 WMIS Application

The major function of WMIS is to provide interfaces for user to input keyword data and find the related information. In WMIS users have to create “personal information needs profile” first, there are two major methods: Input Keyword and Online Help Document. Online Help is txt-format file for supporting Chinese punctuation [5, 7, 10].

When the user profile created their profile then web document-agent can retrieve, analyze and converge web pages for user. The query interface further provides searched results based on request of keyword, author's name, title or publishing date. Furthermore, user can recommend other valuable browsed homepages or provide “Negative Feedback” to WMIS system that can filter non-related homepages. Figure 5 shows results of user query and figure 6 shows advanced query interface in WMIS [2, 3].

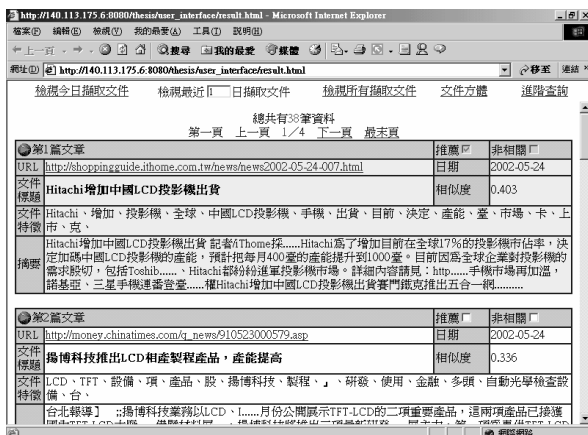


Figure 5 shows results of user query

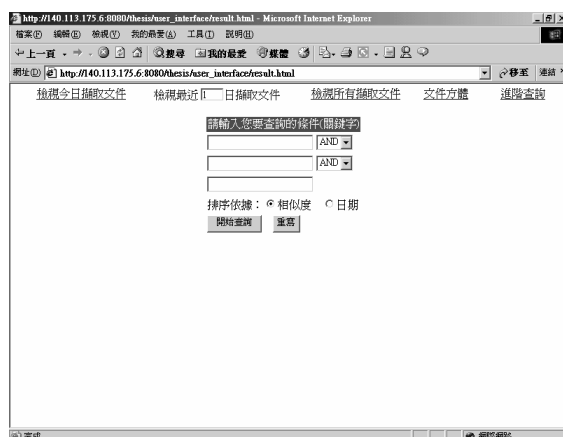


Figure 6 shows advanced query interface

Web mining agent: This is core of WMIS application there are five modules: pre-process document, retrieval web pages, document-related analysis, auto-retrieval abstract and process multi-dimension document. The detailed features of each module are described in the following paragraphs [18].

- **Pre-process document module** includes HTML tag omission, Chinese punctuation and feature retrieval.
- **Retrieval web pages**, this module can retrieve homepages through URL using existing net package of JAVA.
- **Analysis document-related** module uses “Vector Space Model” implement the module through “similarity object”.
- **Auto-retrieval abstract** module can automatically generate an abstract of retrieved document, let users easily understand web pages of contents and save storage space.
- **Process multi-dimension document** module converges every dimensions (including URL, title, publishing date, abstract) store in the document warehouse.

Especially mentioned, the words adopt “Tsai’s List of Chinese Words” of Chih-Hao Tsai¹ at University of

Illinois.

3.4 MyPortal Application

Most companies have an IT environment with disparate legacy systems, applications, processes, and data sources. The system we proposed provides a single interface to access to all structured and unstructured enterprise data. Four design models, business user case model, use case model, analysis model, and design model, are available in MyPortal.

The basic schema of MyPortal is to utilize a web server and applications as the middle tier in a N-tier client/server architecture. With such a schema, it is easily to put existing business model over the Internet. The middle tier is then responsible for accepting the requests for database access initiated at Web browsers, initiating proper database transactions, and returning them to users through the Web server. Based on the idea, the N-tier solution can be designed and maintained more efficiently [15, 16].

4. Discussions

The approach presented in this paper can be enhanced in several aspects: First, The WMIS application can be extended to a multilingual support and enhanced to deal with all kinds of unstructured information. Second, query result cache and paging support can be added to improve data retrieving performance and resource utilization. Third, using middleware component technology, each of these distributed object components can interoperate smoothly in an integrated system. Moreover, there are several object-oriented and component-based development technologies, such as XML, CORBA and EJB (Enterprise JavaBeans), are available. These technologies would be helpful to develop a web-based system, such as WMIS Fourth, SSL, client-side authentication and a single sign-on should be supported in an EIP. Especially, the single sign-on can be integrated with the Windows logon, or it can be extended to support logon to back-end systems after a user has been authenticated to the EIP. A possible approach for achieving the above enhancement is to develop integrated application in the future work.

5. Conclusion

The employee can either use the desktop, notebook, PDA or ad-hoc network devices to access business databases simultaneously and browse the information with a web browser. This paper is a good entry for CIO, one in an organization's IT decision-making team and other interested in IT to understand an EIP solution for organizations. It integrates efficient information of all business intelligent by MyPortal application and provides quickly information search by WMIS application.

The enterprise collaboration portal can be available for every kind of businesses. It also becomes the source of a company’s most valuable asset: the collective knowledge from its employees, suppliers and customers. The “Myportal” prototype is based on the cost-saving

1 Words source: Chih-Hao Tsai
http://casper.beckman.uiuc.edu/~c-tsai4/chinese/tsaiword.zip

conception, it is especially to fit Taiwanese small business. With "Myportal", it becomes possible to overcome the physical boundaries of existing organizations.

Acknowledgements

This research is sponsored by NSC 90 – 2416 – H – 009 – 012.

References

- [1] C. Dias, "Corporate Portals: a literature review of a new concept in Information Management", *International Journal of Information Management*, 21, 2001
- [2] C. L. Yeh and H. J. Lee, "Rule-based Word Identification for Mandarin Chinese Sentences- A Unification Approach", *Computer Processing of Chinese and Oriental Languages*, Vol. 5, No. 2, pp. 97-118, 1991.
- [3] C. S. Peng, S. K. Chen, J. Y. Chung, A. R. Chowdhury, and V. Srinivasan (1998), "Accessing existing business data from the World Wide Web", *IBM System Journal*, 37(1), pp115-132.
- [4] Dan Sullivan, "Document Warehousing and Text Mining", John Wiley, 2001.
- [5] D. Green, "The Evolution of Web Searching", *Online Information Review*, volume 24, number 2, 2000.
- [6] Etzioni, O., and Weld, D. S., "Intelligent Agents on the Internet: Fact, Fiction, and Forecast," *IEEE Expert*, August 1995, pp.44-49.
- [7] G. C. Li, K. Y. Liu and Y. K. Zhang, "Identifying Chinese Word and Processing Different Meaning Structures", *Journal of Chinese Information Processing*, Vol. 2, pp. 45-53, 1988.
- [8] Inmon, W.H., *Building the Data Warehouse*, John Wiley & Sons, Inc., 1993.
- [9] Jun-Wei Yu , "Enriching Enterprise Data Warehouse by Web Mining Technology", Department of Management Science College of Management National Chiao Tung University, Master thesis, June 2002.
- [10] J. Y. Nic and M. Briscobois, "On Chinese Text Retrieval", *Proceeding of SIGIR*, 1996.
- [11] J. Yi and N. Sundaresan, "Metadata Based Web Mining for Relevance", *IEEE* 2000.
- [12] Kuang-Ming Chen , "A Research of Enterprise Information Portal Design", Institute of Information Management College of Management National Chiao Tung University, Master thesis, June 2002.
- [13] M. M. Davydov, "Corporate Portals and e-Business Integration", McGraw-Hill, New York, 2001
- [14] MyYahoo!- personalize Yahoo!'s content, <http://my.yahoo.com>, 2002
- [15] R. Chartier, "Application Architecture: An N-Tier Approach - Part 2", <http://www.15seconds.com/issue/011219.htm>, 2001
- [16] R.R. Ling, D.C. Yen, and D.C. Chou (2000), "From Database to Web Browser: The Solutions to Data Access," *Journal of Computer Information System*, Winter 2000-2001, pp58-63.
- [17] R. Zahavi, "Enterprise Application Integration with CORBA - Component and Web-Based Solutions", John Wiley & Sons, New York, 2000
- [18] San, M., "Intelligent agents on the Internet and Web," *TENCON '98. IEEE Region 10 International Conference on Global Connectivity in Engery, Computer, Communication and Control*, Vol.1, pp.97-102.
- [19] S. Lawrence and L. Giles, "Accessibility and distribution of information on the web", *Nature*, 400, July 1999.
- [20] Ted Sun, "Web-Based Business-to-Employee (B2E) Model: The Strategic and Evolutionary Role in Electronic Business (e-Business) Framework", *Executive MBA Programs Chiao Tung University*, Master thesis, Feb. 2002.
- [21] W. A. Ruh, F. X. Maginnis, W. J. Brown, "Enterprise Application Integration", John Wiley & Sons, New York, 2001