DATA INTEGRATION ISSUES FOR BUSINESS INTELLIGENCE INTEGRATED ENTERPRISE INFORMATION SYSTEMS

Pierre F. Tiako
Center for Information Technology Research, Langston University, Oklahoma, USA

Keywords: Data Integration, Enterprise Information System, Business Intelligence, Distribution, Web Technology.

Abstract Business Intelligence (BI) provides the ability to access any type of data inside or across enterprises and to analyze and present them as usable information. To work on business intelligence, an enterprise has to deal with important problems relating to both (1) Data integration and (2) Analysis and presentation of data for strategic decision-making. No matter what the application, the need for business intelligence applies universally. This position paper focuses on data integration issues for business intelligence integrated Enterprise Information Systems (EIS).

1 INTRODUCTION

Today's enterprises are judged not only on the quality of their products and services, but on how well they share information with customers, suppliers, shippers and banks in the Supply Chain. The more widely available information is throughout the enterprise, the more valuable it becomes. The information actors in the Supply Chain are not in the same location, but are distributed among them. Each actor has its own data representations and presentation mechanisms. This makes sharing and exchanging information across partners extremely difficult. The more integrated an enterprise becomes the easier it is for customers and other partners to access data they need. They are then empowered to make their best decisions to buy or sell.

Several research and industrial contributions (Builders, 1999; Acta, 1999; Enosys, 2002) introduced the concept of business intelligence to refer to an Internet-driven organization that uses information in an intelligent fashion. Business intelligence leverages the ubiquitous Internet network to deploy information consistently and accurately across departments and business units of the same enterprise internally or among autonomous partners externally involved in the Supply Chain. Business intelligence doesn't replace e-business; rather, it is a critical part of an e-business strategy.

This report progress focuses on critical data integration issues for business intelligence integrated EISs. In continuation, this paper is organized as follows. Section 2 is devoted to present concepts of data integration for business intelligence. Section 3 presents critical issues for designing integrated EISs for business intelligence. Section 4 includes discussion on related work and perspectives.

2 BUSINESS INTELLIGENCE AND DATA INTEGRATION

Business intelligence can be regarded as use of information for strategic decision-making. From this point of view, business intelligence mainly deals with data integration and strategic decision-making processes. BI supports the following cognitive phases: (1) Exploration of the world of information, (2) Interrogation of the base of information, (3) Analysis of the base of information, (4) Annotation based on the individual preferences and discoveries (David and Thierry, 2003).

2.1 Business Intelligence

Business intelligence provides the ability to access any type of data inside or across enterprises and to analyze and present them as usable information. To work on business intelligence, an enterprise has to deal with important problems relating to both (1) Data integration and (2) Analysis and presentation of data. No matter what the application, the need for business intelligence applies universally. For *e*-commerce applications targeting the mass consumer market, the most important aspect is intelligence

F. Tiako P. (2005).
DATA INTEGRATION ISSUES FOR BUSINESS INTELLIGENCE INTEGRATED ENTERPRISE INFORMATION SYSTEMS.
In *Proceedings of the Seventh International Conference on Enterprise Information Systems*, pages 366-369
DOI: 10.5220/0002538303660369

about the likes, dislikes, demographics and behavior patterns of customers and prospects; just as critical for business-to-business vendors is analysis about partners, suppliers, and manufacturing and distribution activities. Without business intelligence, competitive *e*-businesses are flying blind. Integration challenge (Kiper, 1994; ETI, 2002) requires different approaches to the problem that supplements the traditional messaging solutions offered by Enterprise Application Integration (EAI). EAIs propose various architectural approaches to integration, but most have a wide range of interfaces and work by simply broadcasting transaction information to other connected systems.

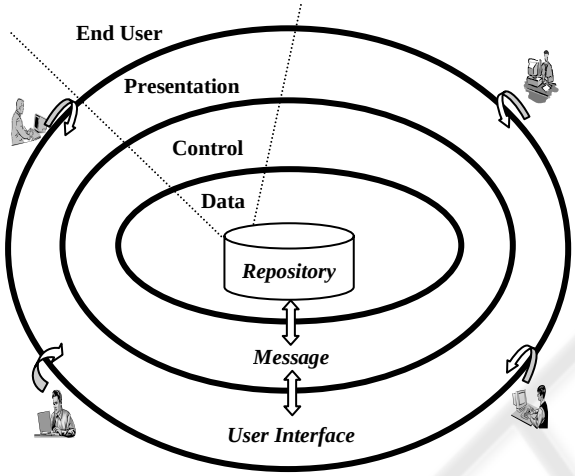

Figure 1: Hierarchical view of an EIS's architecture.

2.2 Hierarchical view of Integrated EISs for Business Intelligence

Each project related to tools integration (Brown, 1992) must be categorized according to the five dimensions of integration proposed by Wasserman (1990), Thomas and Nejmeh (1992). The dimensions are presentation, platform, control data, and process integration. For software environments resulting from integrated project support environment projects, Brown (1992) considers three aspects of integration (see figure 1): presentation, control and data integration. Presentation integration deals with the way users interact with EISs using services of user interface manager. Control integration aims at coordinating tools execution (Brown and Wallnau, 1996). A tool invokes another for realizing an action, rather than implementing that action. Data integration deals with the problems of sharing or exchanging data between tools.

2.3 Research and Industrial Approaches

Industrial solutions dealing with data integration favorable to business intelligence (see table 1) can be grouped into the following categories: (1) EAI systems, which are vital to the creation of collaborative e-business applications because they enable the integrity of data and business rules to be preserved; (2) Business-to-business integration (B2Bi) systems that focus on extending EAI between companies across the Internet (they typically do not offer in-depth data integration capabilities beyond Web-scraping technologies); and (3) Enterprise Information Integration (EII) that can be distinguished from the other by its underlying technology that provides real-time, seamless access to any data, anywhere, for any application.

Depending on each enterprise's unique situation, EII can be applied to perform rapid, high-level data integration either as a stand alone solution or within the framework of a larger integration project. Because EII technologies work over standard web connections to access an array of data sources (Dossick and Kaiser, 1996) they can be invoked as a web service or embedded within an integration application rapidly and easily.

Table 1: Existing Technologies for integrated

Category	Definition	Function	Technology
(Industry) Enterprise Application Integration (EAI) Systems	Systems that link disparate enterprise applications	Integrate applications at business process level with focus on transaction.	RMI Applet (Java Soft) OLE (Microsoft) ODBC JDBC APIs
(Research) Control Integration (SoftBench, 2002) Field(Reiss, 1994)	Approaches that control tool during integration process	Handle integration messages exchanges between EIS	CORBA (OMG) DCE (OSF) SOM/DS0M (IBM) (SOM, 1994)
(Industry) Business-to-business integration (B2Bi) Systems	Systems that automate point-to-point trading transactions among partners	Help migrate to the web without disrupting existing business practices	HTML CORBA (OMG)
(Research) Data Integration (Boudjlida,1995)	Contributions that deal with data integration during the integration process	Support semantic and syntactic information relating to data during integration	CORBA (OMG) DCE (OSF) SOM/DS0M (IBM)
(Industry) Enterprise Information Integration (EII) Systems	Systems that access, integrate, and unite data from multiple sources	Provide real-time data access within and through corporate firewalls.	XML CORBA (OMG)

The same research contribution dealing with data integration in favor of business intelligence can be group into the following categories: (1) data integration, which is concerned with data representation and naming. Significant contribution carried out in data integration approaches proposes a dynamic approach based on an abstract model, which captures data semantics and constraints. The model includes a set of constructors, which allow defining complex structures and a set of operators that determine the transformations that could be applied to change one type to another. (2) Control integration, which is concerned with logic or structure of programming model. Control also implies that a tool can decide whether and how much to share operations it supports and data it manages. Thus a tool could be added or removed from an EIS; it could participate in more than one interoperating EIS.

Despite the number of research and industrial approaches to address, they are almost partial because each of them is expressively aimed at addressing a particular aspect of integration. Again no matter what the application, the need for business intelligence applies universally. The idea of data integration for business intelligence is to provide a more general architecture supporting tool integration, for we need to combine a set of mechanisms in developing a complete and coherent system, as presented below.

3 CRITICAL ISSUES FOR DESIGNING INTEGRATED EII FOR BUSINESS INTELLIGENCE

The critical issues for designing Integrated Enterprise Information Systems for Business Intelligence (IEISBI) are listed below. IEISBI have to address these issues in order to be competitive.

Recent research, as well as new products, focuses on Web Services as a means to integrate data. However, Web Services are new technology and all the requirements discussed above and issues listed below still apply. For instance, semantic Web Services (WSMO, 2004) addresses the requirements and issues taking all aspects into consideration.

(a) Communication. IEISBI technology must provide reliable and secure communication between back end applications in order to ensure overall consistency

(b) Distribution. IEISBI applications are distributed in the sense that each has its own separate storage or database management systems with each separately controlled and managed

(c) Heterogeneity. IEISBI applications are implemented based on their own data model and due to the particular management focus the data models differ in general not complying with a common standard.

(d) Mediation. Due to heterogeneity of IEISBI applications, objects are represented in different data models that have to be mediated if objects are sent between applications.

(e) Process management. Multi-step business processes across IEISBI applications require process management (Tiako, 1998) to implement the particular invocation sequence logic.

(f) Reliability. The integration of IEISBI applications must be reliable in order to neither loose data nor accidentally introduce data by retry logic in case of failures

(g) Replication. Replication is a mechanism that ensures that changes of an object are automatically propagated to duplicates of this object. The various representations act as if they are one single representation

(h) Security. Communication of data between IEISBI applications through EIS technology must ensure security to avoid improper access.

4 RELATED WORK AND PERSPECTIVES

Here we briefly present some EIS tools available in the market and susceptible to favor business intelligence by data integration. WebFOCUS (Builders, 1999) connects core business systems, enabling data to flow freely from one to another. ETI (2002) automate the process of data collection, transformation, and migration between incompatible systems in heterogeneous computing EISs. ActaWorks (Acta, 1999) solutions connect ERP systems to business to business web applications, to business to commerce web applications and to direct-to-the-web data access applications. Informatica (2000) outlines the current state of convergence of e-business and business intelligence, and shows how its data integration solutions support the key infrastructure demands of e-businesses. Enosys (2002) shows how XML presents significant advantages as the data representation language for modern *e*-business applications and addresses their needs better.

The emerging of the Internet has accentuated the need for supporting a broad diffusion of information and the openness of EISs. To meet these needs, user-friendly interfaces must be based on HTTP protocol (Riva and Ramoni, 1996) and HTML. So, user interfaces should benefit from standards of scripting programming languages such as JavaScript (Flanagan, 2002) for increasing the interactivity with users.

The distribution of components of an EIS can be realized by coupling Web technologies and distributed object approaches (Brose et al., 2002) such as CORBA or COM, including a Java implementation of its specifications. The resulting platform built on top of this kind of association are opened, and characterized by the following properties:

(a) **Neutrality** of the platform. Means independent of the underlying operating system and hardware architecture.

(b) **Portability** of components. Components of an EIS can be moved from one platform to another, with little effort, by adapting the new component for its integration in an already existing EIS.

(c) **Distribution**. The user interface managed by the Web browser is completely separated from the server that supports the processing and the storage of data. Interoperable tools which make up an EIS are installed on different computers.

This way of distributing components of an EIS, by using the Web as support, is adopted by different approaches, which are grouped here in two categories. The difference between these proposals lies in methods the Web accesses legacy server applications and their data. Data integration issue for business intelligence presented in this paper differs from the above solutions by its openness. The future of this work will study the issues presented above.

REFERENCES

David, A., and O., Thierry, 2004. *EQuA2te Architecture and its application in Business Intelligence*. Business Intelligence: Research & Applications Conference - IERA 2003 - Nancy, France.

Acta, 1999. *ActaWorks™ RealTime Technical Overview: A Data Integration Platform For eCommerce And Business Intelligence*. A White Paper - Acta Technology. http://www.acta.com.

Boudjlida, N., and O., Perrin, 1995. *An Open Approach for Data Integration*. In 2nd Int. Conf. on Object-Oriented Information Systems (OOIS'95), Dublin.

Brose, G., et al., 2002. *JacORB Programming Guide 1.4.1*. Freie Universität Berlin and Xtradyne Technologies AG.

Brown, A., W., 1992. *Control Integration through Message Passing in a Software Development Environment*. Technical Report, CMU/SEI-92-TR-35.

Brown, A., W., and K., C., Wallnau, 1996. Engineering of Component-Based Systems. *Component-Based Software Engineering: Selected Papers from the Software Engineering Institute*. IEEE CS Press.

Builders, 1999. *The Critical Role of Business Intelligence in e-Business*. A White Paper - Information Builders

Dossick, S., E., and G., E., Kaiser, 1996. WWW Access to Legacy Client/Server Applications. *Computer Networks and ISDN Systems, The International Journal of Computer and Telecommunications Networking*, 28(7-11):931-940, Elsevier Science B.V.

Enosys, 2002. *Enterprise Information Integration: Integrating the Real-time Enterprise*. Business White Paper 8 Enosys Software Confidential.

ETI, 2002:. *Data integration management for the enterprise*. The ETI solution TM version 5. Evolutionary Technologies International. http://www.eti.com.

Flanagan D., 2002. *JavaScript: The Definitive Guide*, 4th Edition. Sebastopol, Ca.: O'Reilly.

Informatica, 2000. *Business Intelligence and the New World of E-business*. A White Paper WP1049-B. Informatica. www.informatica.com

Kiper, J., D., 1994. A Framework for Characterization of the Degree of Integration of Software Tools. *Journal of System Integration*, No. 4, pp. 5-32.

Reiss, S., P., 1994. *The FIELD Programming Environment*. Kluwer Academic Publishers.

Riva, A., and M., Ramoni, 1996. LispWeb: a specialized HTTP server for distributed AI applications. *Computer Networks and ISDN Systems*, 28(711):953-961, 1996.

SoftBench, 2002. *Integrating with HP SoftBench*. HP-UX STK version 1.8.1. Hewlett-Packard.

SOM, 1994. *The System Object Model (SOM) and the Component Object Model (COM):* A comparison of technologies from a developer's perspective. IBM Corporation, Object Technology Products Group

Thomas, I., and B., A., Nejmeh, 1992. Definitions of Tool Integration for environments. *IEEE Software* Vol. 9, N° 3, pp 29-35.

Tiako, P., F., 1998. Modeling the Federation of Process Sensitive Engineering Environments: Basics Concepts and Perspectives. In *Software Process Technology* (V. Gruhn, Eds) - LNCS Vol. 1487, Springer-Verlag.

Wasserman, A., I., 1990. Tool Integration in Software Engineering Environments. In *Software Engineering Environments. International Workshop on Environments*, LNCS N° 467, pp.137-149.

WSMO, 2004. *Web Service Modeling Language*, www.wsmo.org.