Risk Assessment Of Data Warehousing, A Case Study

Bongsik Shin
Information and Decision Systems
College of Business Administration
San Diego State University
E-mail: bshin@mail.sdsu.edu

Abstract

Risks are factors frequently beyond control and can negatively impact organizational ability to achieve the established goals. These are factors that create differences between the reality and an ideal situation. Accurate evaluation of potential risks is integral to a large-scale project, especially when the project has enormous and long-term organizational implications. In that aspect, the risk evaluation for data warehousing is a *readiness* test to measure preparedness of an organization for the massive project. Although data warehousing concept is rapidly spreading in , many projects have ended up in failure due to the lack of organizational readiness and incomplete risk management. Partial understanding of the risks inherent to data warehousing can be made from the theoretical perspective of critical success factors (CSF) of information systems. However, because of the architectural and procedural complexity, risk assessment of data warehousing demands more comprehensive estimation of other issues as well. A case study was undertaken on the risk management of a data warehousing project. The research focus was placed on identifying main risk factors recognized by the project team in undertaking the large-scale effort. Potential threats to the project were classified under the categories of external dependencies-, organization-, planning-, business case-, and technology-relationship. Technological risks were again divided into environmental, project-related, and operational ones. The intensity of each identified risk was then represented via a quantitative indicator. The process resulted in a comprehensive understanding of potential sources of trouble and the evaluation of preventive measures that could be taken.

1. Introduction

Data warehousing has become the central point of information management and the integrated backbone of decision support systems that contain a large volume of historical data. From being a little-understood experimental concept years ago, it has recently reached a stage at which nobody is questioning its strategic value. Statistical indicators and surveys indicate that the number of companies which either already own or currently building a data warehouse is exploding and large enterprises are involved in at least one or more data warehouse projects (Hyperion, 1999; Sen and Jacob, 1998). Because of its architectural complexity, enormous implementation cost, and difficulty to demonstrate return on investment (ROI), the data-warehousing concept has been somewhat controversial and sometimes encountered managerial opposition (Watson and Haley, 1998). However, its value is well justified by the fact that databases tuned for operational and transactional use are, in general, not structured to effectively satisfy the information demand from decision-makers.

Various motivations for data warehousing have been reported, most noticeably the industry-wide emergence of stiff competition that forces organizations to adopt innovative measures to increase organizational competency. Handen and Boyle (1998) indicated that deregulation, globalization, and partnering are major forces driving adoption of data warehousing in order to meet intensified competition in the telecommunication industry. Growing popularity of distributive computing and virtual processes such as virtual corporations, telework, mobile work, and EDI (electronic data interchanges) may fuel the growth of integrative data warehousing to offer better information access to constituencies.

Other motivational factors are the rapid growth of database volume, requirements for reliable system availability, reduced operating expenses, improved customer service, increased performance, flexibility, and stability in accessing corporate information. Organizations also may be looking for appropriate business process re-engineering to take advantage of the single source of data that data warehousing offers (Watson and Haley, 1998). For instance, consolidation and streamlining of information generation by enforcing organization-wide consistency in data architecture and related metadata reduce unnecessary duplication of knowledge work.

Despite the enormous impact this technology will have on business performance, academic research has been rare. The motivation for this case study was to investigate the risk assessment approach taken by the data warehousing project team at one of the largest insurance companies in the US. Because of the architectural and procedural complexities and large financial commitment, a data warehousing project demands comprehensive estimation of potential risk factors prior to its launching. This research focus was placed on discussing main risk factors recognized by the project team in undertaking the large-scale effort.

2. Motivation

The company studied is a large Fortune 500 enterprise that provides individual health and accident insurance services. It covers more than 6 million individuals with over $80 billion of life insurance in force and has more than 8,000 employees. The company maintains a national network of 150 sales offices and a sales staff of more than 3,100. Through its several affiliates, it also provides a variety of other services including mutual fund investment, full-service brokerage, investment banking, foreign currency exchanges, travel insurance, specialty life plans, and health and accident plans.

The company developed a long-term strategic plan to effectively align information technology with business goals, and chose data warehousing as the initial project. First, it recognized that improvements in data management and delivery process were needed to better serve its long-term business strategy. In the insurance service industry, effective management of service data and its utilization is a critical business success factor. Executives in the company determined that it needed to refine data both to manage information better and to improve its internal business information strategy.

Like any large enterprises, the company maintained a number of operational and transactional databases. Effective integration of these data sources was becoming an imperative because they have evolved independently resulting in much data duplication, inconsistency and other data quality problems. The company had created operational data stores (ODS)[1] that combined operational data on a common subject from transactional systems. Corporate ODS such as the financial system had a broader scope that could be accessed across the company whereas such a departmental ODS as the client management system was limited to departmental applications. The ODS were using data warehousing technology to provide a holistic view of data for operational purposes. It therefore supported day-to-day business processes of the company and was not used to support analytical processing. Naturally, in the absence of an integrated environment for decision support applications, corporate executives considered data warehousing critical infrastructure.

Additionally, data warehousing was regarded crucial to curtailing duplicative efforts by business units to obtain the same kind of information from different databases. Business units of the company were organized into three different customer levels: individual, group, and corporate. Despite the customer-orientation emphasis on organizational structure, the information requirements from the individual customer and the group customer business units overlapped considerably. In a traditional non-integrated database environment, duplicative efforts to get the same information were difficult to avoid. Furthermore, reports produced by three groups were inconsistent because of the heterogeneity in the data acquisition and their processing. The company understandably wanted to establish a central source of decision support data. This environment would not only streamline the information production procedure, also it could minimize procedural wastes and deliver reports in a consistent manner. In fact, top management thought that substantial reduction (or possible elimination) of the process overlap was an enough justification for the multi-million dollar investment.

3. Risk Assessment

Potential risks associated with the data-warehousing project were identified and evaluated prior to its formal launching. Risks are factors frequently beyond the control and have negative impacts on the ability to achieve the established goals. Many risk factors in information system projects have been discussed, including the size and structure of information systems, the skills of system analysts, and the familiarity with technology (McFarlan and

[1] ODS was defined as a “subject oriented, integrated, volatile (that is up-datable), current valued data store containing only corporate detailed data (Inmon, 1996).

McKenny, 1983; Saarinen and Vepsalainen, 1993). Risk evaluation is integral to a large-scale project (Vitale, 1986), especially when the project has enormous and long-term organizational implications. Even with the significant payback estimated at 400 percent in 3 years and the large investment of capitals averaging at $3 million, many data warehousing efforts have ended up in failure due to the lack of organizational readiness and the incomplete risk management (Murtaza, 1998). It was estimated that the chance of data warehousing failure in the initial stage is as high as 50 percent (Watson and Haley, 1998). Risk evaluation at the company was therefore conducted as a *readiness* test to assess whether it was prepared for the successful undertaking of the data warehousing project.

Furthermore, as the project was decided to be undertaken by in-house development staff who were relatively inexperienced in data warehousing, estimating project risks and determining problem resolution strategies were crucial not only to justify the project launching but also to curtail potential threats of failure once it was started. Risk management can be more systematic when relevant risk factors and their resolution approaches are determined *a priori* based on a certain taxonomy scheme. For instance, Baskerville and Stage (1996) suggested that risks are associated with systems developers, users, application domain, problem domain, computer system, and development environment and suggested risk mitigation through a cycle comprised of detailed risk definition, specification of consequences, and selection of resolution strategy. Saarinen and Vepsalainen (1993) generalized risks in terms of system complexity and uncertainty.

The project team classified potential threats into external dependencies-, organization-, planning-, business case-, and technology-related ones (Table 1). Technological risks were again divided into environmental, project-related, and operational categories. Severity of each identified risk was then evaluated and represented via a quantitative indicator. This approach enabled risk assessment of the data warehousing project at both micro and macro levels.

Overall, the itemized weighting indicated moderate risks in planning, organizational and technological aspects. Technological risk was regarded higher than planning and organizational risks, mostly because the project was based on unfamiliar technologies and involved a number of physical-system interfaces among various heterogeneous components. Furthermore, user applications to be adopted on the data warehouse had not yet been clearly defined. Lack of packaged solutions that reduce technical uncertainties and speed up the development process, and expected high-volume system throughput were also pointed out as the main technological risks of data warehousing.

Organizational risks were regarded relatively weak because, above all, the top management was highly supportive of data warehousing and rendered a necessary sponsorship. However, the project team identified several potential risk items in this category. First, it was suggested that the absence of clearly defined key users of the new system could pose a potential threat. Appropriate feedback from target business users is critical to a successful data warehousing effort (Mutaza, 1998). Naturally, when key users of the system are not clearly identified and their requirements are not effectively incorporated, the final product may not be widely accepted by prospective users.

Second, changes of organizational and user procedures needed by the new system could possibly create a roadblock to the success of the project. In fact, it was pointed out that anticipated benefits of data warehousing may be difficult to attain unless necessary organizational and user changes occur as a form of redesign in business processes (Watson and Haley, 1998). Possible resistance to or rejection of organizational changes incurred by the innovation has been well documented theoretically and empirically (Abrahamson, 1991; Damanpour, 1991). Besides, the project team appeared to be concerned with that the control procedures of the data warehousing project had not been well defined and therefore remained rather informal. The lack of formal controlling mechanisms for the project management appeared to be rather natural because the development staff had little experience in data warehousing. Naturally, the overall project entailed much uncertainty in terms of the project management, which includes project planning, change management, user-expectation management, scope management, and communication planning.

Risks caused by external dependencies and the business case were expected to be relatively low because the project was to be undertaken by the in-house staff. The risk of overlapping between the data warehousing project and other existing activities was pointed out. For instance the IS department was directly or indirectly involved in multiple activities including a financial data mart, operational data stores, and a comprehensive Y2K projects. Such an overlapping could incur problems in resource mobilization, end-user support, and system redundancy and leverage. In particular, with the complexity, data warehousing efforts demand dedicated human resources (ex., programmers, project managers, database modelers, support, etc). The resource constraint was especially keen at that time because the project members were assigned to other projects as well. As for the business case, the main risk was related to the

fact that the project was initiated to improve the business efficacy of the company and therefore was fundamental to its long-term business strategy. Its failure, therefore, could considerably disrupt the company's phased IT plans and subsequently affect business operations.

Attempts were made to minimize identified risks whenever possible by introducing mitigation factors although eliminating all risks was not feasible. For instance, there was some overlapping between the planned data mart pilot project and on-going activities (ex., departmental operational data stored) as a measure to provide improved access to *Compliance Monitoring* data. *Compliance Monitoring* was an application to monitor, analyze, and report the regulations compliance of insurance agents and brokers who sell products and services for the company. To reduce the resulting confusion and inconsistency between two overlapping activities, some members (board members and analysts) from existing projects were also assigned to the data mart project on a part-time basis. Overall, although risks are contingent on organizational circumstance, identified risk items in this case study offered valuable insights on potential threats that could inhibit the successful planning and deployment of corporate data warehousing.

Risk categories	Significant risk factors identified
External dependency	1. *Overlapping scope with other developments* 2. *Plan requires extensive recruitment of resources.*
Organizational risks	1. *Number of user areas/decision makers* 2. *Key users unavailable* 3. *Organizational changes required* 4. *Level of changes required to user procedures*
Planning risks	1. *Critical implementation date* 2. *Informal control procedures* 3. *Effort development years vs. elapsed time* 4. *Planned resources available*
Business case risks	1. *Fundamental to IS strategy*
Technological risks	Environmental: 1. *New/unfamiliar technology* 2. *Use of development method/standards* Project-related: 3. *Complexity of functions* 4. *Complexity of database* 5. *Shared databases with undefined applications* 6. *Number of physical systems interfaces* 7. *Package solution available* 8. *Multi-level hardware implementation* Operational: 9. *High-volume throughput*

Table 1. Summary of key risk factors identified

4. Concluding Remarks

Numerous factors affect the fate of the data-warehousing success. Some of these are more generic in the sense that their influences can be generalized regardless of the types of information systems. These variables have been frequently studied under the umbrella concept of critical success factors (CSFs) of information systems. Strong support from business management, proper change management, effective upward and downward communication, adequate alignment between business strategy and information technology design, appropriate rewarding structure for system designers and users, user-centered design, user involvement, user training and education are among the many.

While aforementioned are general success factors of information systems, some factors are believed to be rather unique to the data warehousing environment (Berg and Heagele, 1997; Edholm, 1997; Sigal, 1998). Above all, much of the uniqueness results from the fact that data warehousing entails a stiff learning curve because of the system complexity in planning, design, and implementation. Naturally, the complexity emphasizes crucial importance of comprehensive risk assessment of a data warehousing project prior to its initiation. Although the risk items were subjective and tended to be organization-context binding, this case study offered some understanding of potential sources of trouble in data warehousing efforts and showed the importance of preventive measures to avoid them.

References

[1] Abrahamson, E., "Managerial Fads and Fashions: The Diffusion and Rejection of Innovations," *Academy of Management Review* 16(3), 1991, 586-612.

[2] Baskerville, R. L. and Stage, J., "Controlling prototype development through risk analysis," *MIS Quarterly*, December 1996, pp. 481-554.

[3] Berg, D. and Heagele, C., "Improving data quality: A management perspective and model," in *Building, Using, and Managing the Data Warehouse* edited by Barquin, R & Edelstein, H., Prentice Hall, 1997.

[4] Damanpour, F., "Organizational innovation: A meta-analysis of effect of determinants and moderators," *Academy of Management Journal* 34(3), 1991, 355-590.

[5] Edholm, K.G., "Bring user perspective to data warehouses: Twenty-one points of consideration," in *Building, Using, and Managing the Data Warehouse* edited by Barquin, R & Edelstein, H., Prentice Hall, 1997.

[6] Handen, L. and Boyle, P., "Data warehousing: managing knowledge," *Telecommunications*, vol. 32, no. 1, 1998, pp. 61-68.

[7] Hyperion Co., "The role of the OLAP server in a data warehousing solution," White Paper <http://www.hyperion.com/whitepapers.cfm>May 13, 1999.

[8] Inmon, W. H., *Building the Data Warehouse*, 2nd Edition, ed. John Wiley, 1996.

[9] McFarlan, F. W. and McKenny, J. L., "*Corporate Information Systems Management – The Issues Facing Senior Management*. Irwin, Homewood, Illinois. 1983.

[10] Murtaza, A. H., "A framework for developing enterprise data warehouses," *Information Systems Management*, fall 1998, pp. 21-26.

[11] Saarinen, T. and Vepsalainen, A., "Managing the risks of information systems implementation," *European Journals of Information Systems* 2(4), 1993, 283-295.

[12] Sen, A. and Jacob, V. S., "Industrial-strength data warehousing," *Communications of the ACM*, vol. 41, no. 9, 1998, pp. 28-31.

[13] Sigal, M., "A common sense of development strategy," *Communications of the ACM*, vol. 41, no. 9, 1998, pp. 42-43.

[14] Vitale, M.R., "The growing risks of information systems success," *MIS Quarterly*, Dec. 1986, pp. 327-334.

[15] Watson, H. J. and Haley, B. J., "Managerial considerations," *Communications of the ACM*, vol. 41, no. 9, 1998, pp. 32-37.