

December 2002

APPROACHES TO DATA WAREHOUSE DEVELOPMENT: CONTINGENCIES AND APPROACH SUCCESS EVALUATION

Thilini Ariyachandra
University of Georgia

Follow this and additional works at: <http://aisel.aisnet.org/amcis2002>

Recommended Citation

Ariyachandra, Thilini, "APPROACHES TO DATA WAREHOUSE DEVELOPMENT: CONTINGENCIES AND APPROACH SUCCESS EVALUATION" (2002). *AMCIS 2002 Proceedings*. 341.
<http://aisel.aisnet.org/amcis2002/341>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2002 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

APPROACHES TO DATA WAREHOUSE DEVELOPMENT: CONTINGENCIES AND APPROACH SUCCESS EVALUATION

Thilini Ariyachandra
University of Georgia
thilinia@uga.edu

Abstract

Currently there is much debate in industry regarding which data warehouse development approach best fits a given organization setting. Using organizational information processing theory, this research proposes a model that identifies the various rational contingency factors that affect the choice of data warehouse development approaches. It further describes social political contingency factors that could impact the development approach selection decision. Consequently, the research also identifies means of evaluating the success of the various approaches. An empirical investigation involving interviews with experts and a field survey is proposed to further validate and test the proposed model. In so doing, this research intends to shed light on the data warehouse development process in organizations and evaluate the impact of rational vs. social political factors on development approach selection and success.

Keywords: Data warehousing, systems development approaches

Introduction

From its introduction as another IT solution for data integration issues in organizations over a decade ago, data warehousing has evolved to being recognized as a core requirement in most organizations' business intelligence (BI) environment. It is being recognized as a mandatory component of the critical infrastructure required for strategic organizational initiatives like customer relationship management, performance management and supply chain integration. A 2001 study conducted by Survey.com revealed that the data warehousing market is set for steady growth in the next three years, despite signs of the global economy slowing down. The study indicated that organizations would expand their current data warehousing and BI annual spending by 74 percent in the next three years.

Despite growing popularity and importance, there is no industry consensus on the best warehouse development approach to use. Currently, there are several competing approaches: top-down, bottom-up, independent data mart; and federated. This has left novice data warehouse developers with no clear direction on how to embark on this high cost investment. Consequently, the lack of a specific, clear methodology to guide development efforts has been reported as the third most severe problem encountered in developing data warehousing applications by data warehouse system designers (Griffin 1997).

No approach has been associated with a 100 percent success or failure rate. Thus, contrary to the claims of experts of the different warehouse development approach camps, there is no one approach that succeeds above the rest. According to Joshi (1999), the best development approach for IS systems is dependent on the circumstances of the organization. Different contingency factors related to the specific requirements and situation of the organization determine the information systems development approach used (van Slooten and Schoonhoven 1996). Given the growing demand for data warehouses, the tremendous dollar investment required for their development, and the high rate of initial failures, there is a need to (1) discover if and what contingency factors affect data warehouse development approaches and (2) evaluate which of the dominant data warehouse development approaches best fit an organizational setting. The current study proposes to satisfy these needs and provide both practitioners as well as

academics a better comprehension of the development approach selection process as well as an evaluation of the success of the various approaches.

Research Questions

In order to better understand the factors that affect the selection of a data warehousing development approach and the success of various development approaches, the following research questions will be investigated:

1. What factors influence the choice of data warehouse development approach?
2. How effective are the alternative data warehouse development approaches?

Research Model

Having decided to build a data warehouse, the selection of a data warehouse development approach is another one of the many decisions faced by organizations. This warrants the application of decision-making theories to the development approach selection process. Literature on organizational decision making describes two competing theories for effective decision making in organizations (March and Simon 1958; Cyert and March 1967; Pfeffer 1981). These competing theories are rooted in the rational analytical and social political schools of organizational theory. The theories in each school of thought describe contingencies that affect the organizational decision-making process. Contingency factors are circumstances of the project that in some way influence the selection or construction of an approach to systems development (van Slooten and Schoonhoven 1996).

Contingency Factors

The proponents of the rational decision school argue for the use of comprehensive information processing to attain an organizational outcome-maximizing goal (Sambamurthy, Venkatraman et al. 1993; Daft 1998). Organizational information processing theories, part of the rational analytical school of thought, provide some insight into the decision-making process involved in choosing the best data warehouse development approach (Thompson 1967; Galbraith 1973). The selection of a development approach involves an assessment of the information processing requirements and the information processing capacity of the organization.

The information processing requirements of the organization can be defined in terms of the complexity of the tasks that need to be accomplished using the system, the nature of the task environment, as well as the unit interdependence within the organization (Tushman and Nadler 1978; Goodhue, Wybo et al. 1992). In a data warehouse development approach selection context, the complexity of the task can be described in terms of the need for predefined vs. ad hoc query support to accomplish tasks. The nature of the task environment can be described by the nature of the dynamic turbulent environment surrounding the organization and the unit interdependence by the variations in organizational structure. The information processing capacities can be described in terms of existing IT capacities (i.e., the nature of existing legacy systems) and existing data standards (i.e., the quality of legacy systems in terms of data readiness for integration). Together, the requirements and the capacities form the characteristics of the organizational context relevant to the selection of a development approach.

The social political school views decision making as a process of negotiation and coalition building in which multiple ambiguous goals exist (Pfeffer 1981; Markus 1983). This view questions the assumption that an overarching organizational goal accomplishment objective influences decision making. In the context of data warehousing, it challenges the notion that an overall company goal determines the development approach selection decision. The presence and impact of multiple organizational goals as well as the political decision-making process described in this school of thought, can be described in terms of contingencies faced by potential data warehousing organizations such as: nature of current data ownership, the importance and credibility of the IS function in the organization, the source of sponsorship and degree of sponsor commitment in the organization.

Past research has applied both rational analytical theories and social political theories to the organizational decision-making process. No one set of theories has explained more of a situation than the other. Consequently, more recent literature calls for the application of both competing theories to explain organizational decision making (Ribbers, Peterson et al. 2002). A survey of the practitioner literature related to data warehouse development approaches further validates the inclusion of contingencies related

to both schools of thought. Upon assessment of these contingencies (i.e., characteristics) associated with the organizational situation, the development approach that provides the best fit between the requirements and the capacities can be chosen.

The fact that contingencies relevant to the context determine the best approach to use is strongly supported by the cognitive fit theory (Agarwal 1996). Cognitive fit theory states that problem solving elements should fit the characteristics of the system design task. This theory can be applied to systems development as follows: the characteristics of the methodology being used should match the characteristics of the application to which it is being applied (Howard 1999). In the context of data warehouse development, this means that the characteristics of the development approach used should match the characteristics of the organizational situation in which a data warehouse is being built.

Developmental Approaches

At present there are two main competing approaches being advocated for data warehouse development: top-down and bottom-up. In addition, independent data mart and the federated architecture approach are two other approaches that are also applied for data warehouse development today.

In the top-down approach, an enterprise data warehouse is built in an iterative manner, business area by business area, and dependent data marts are created as required from the enterprise data warehouse. This approach leads to the gradual creation of a centralized data repository with common data standards and integrated data elements for the entire organization.

In the bottom up approach, data marts are created with the view of integrating them into an enterprise data warehouse in the future. This approach is characterized by the step-by-step development of subject area data marts that fits an overall enterprise wide architecture. The first mart is built as a proof of concept and additional marts are added according to a master suite of conformed dimensions and standardized fact definitions.

The independent data mart approach, or stovepipe approach, is similar to the bottom up approach except that it has no overall enterprise wide architecture, conformed dimensions or fact tables. Each data mart is built independently of the others, often involving different platforms and different data definitions. Finally, the federated architecture approach involves integrating existing data warehouses and data marts in the organization by establishing shared, common data points across all systems. Unlike the top-down approach, the federated approach allows for dedicated data feeds for each individual system and does not have a single central data repository.

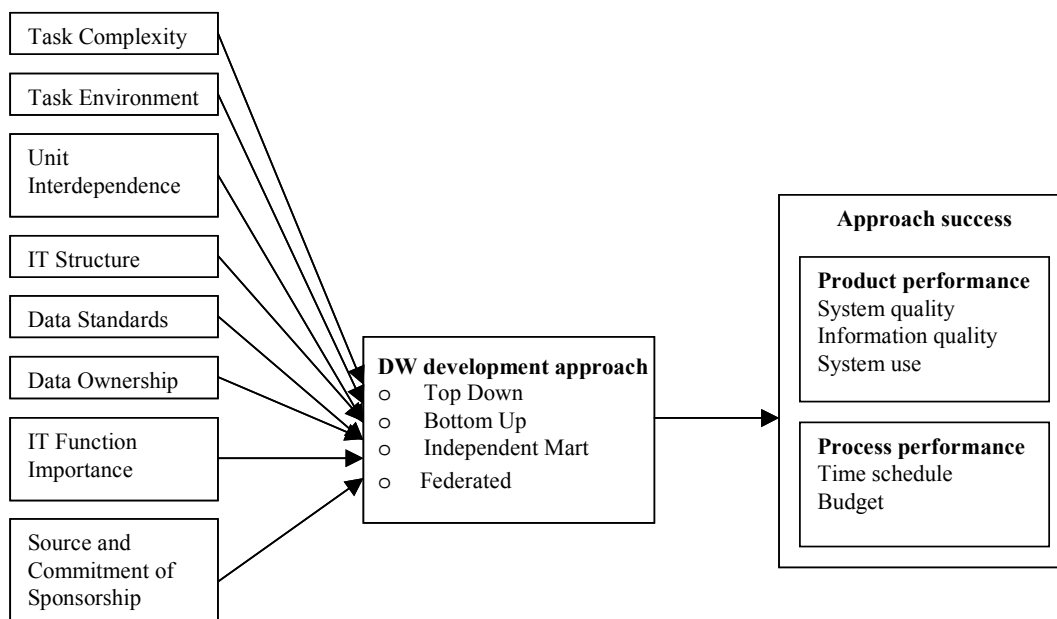


Figure 1. The Proposed Research Model

Approach Success

In order to evaluate the fit of the developmental approach to the organizational context, a measure of approach success needs to be assessed (Figure 1 contains the proposed research model). Approach success can be described as the efficiency and effectiveness of the data warehouse developmental approach used. It takes into account two key dimensions: process performance (how well the process of data warehouse development and the project went) and product performance (how quality of the developed system, that is, the product or output of that process, is). Each dimension needs separate assessment, as they are not necessarily highly correlated. For example, it is quite possible for an over-budget or beyond-schedule project to deliver a high-quality product. Conversely, a within-budget and on-time project may deliver a product of poor quality.

The product performance dimension draws parallels to the information systems success model developed by Delone and Mclean (1992; 2002). Of the six main dependent variables described by Delone and Mclean (1992; 2002), system quality, information quality, and system use constructs have been used to assess IS success in previous IS project management and project success assessment research. In the context of development approach evaluation, process performance can be assessed in terms of time schedule and budgetary considerations for the overall project.

Research Methodology

This study involves two phases. The initial phase entails interviewing a panel of experts to further enrich existing literature on contingency factors and success measures described in the research model. The expert interviews and existing literature will lead to the creation of a comprehensive exploratory model that describes the most suitable approach for different contingency combination sets.

In the second phase, the main task of study, will involve conducting a field survey of data warehousing organizations, targeting data warehousing managers as survey respondents, to gain empirical support for the exploratory model created in phase one. The unit of analysis for the study will be the data warehouse development approach.

After survey questionnaire validation, pilot testing, and data collection, the analysis of data will involve a series of statistical analyses. First, the overall measurement model will be assessed using structural equation modeling techniques to establish validity and reliability of measures. Next, the data collected on the different contingency factors and the actual approach used, will lead to the development of a predictive model of approach used for a given combination of contingency factors. This analysis will be extended to compare the fit (e.g., the differences) between the actual development approach used and the ideal approach proposed by the exploratory model.

Next, approach success measure calculations will be used as the basis to determine the ideal approach for a given combination of contingencies. Consequently, the measures will also help determine the degree of variation between the empirically calculated ideal approach and the theoretically determined ideal approach, proposed by the exploratory model, for a given contingency combination. This may lead to the identification of new relationships between contingency factors and approach used as well as how these relationships impact approach success.

Consequently, the overall study is expected to provide insight into the contingent nature of the data warehouse development approach. Furthermore, it would help practitioners select a development approach that is appropriate for their organizational situation. Overall, the study is intended to give both practitioners as well as academics a better understanding of the data warehouse development process in organizations.

References

- Agarwal, R. "Cognitive fit in requirements modeling: A study of object and process methodologies." *Journal of Management Information Systems* (13:2), 1996, 137-163.
- Cyert, R. M. and J. G. March. *A behavioral theory of the firm*. Englewood Cliffs, New Jersey, Prentice Hall, 1967.
- Daft, R. L. *Organization theory and design*, South Western College Publishing, 1998.
- Delone, W. H. and E. R. Mclean. "Information systems success: quest for the dependent variable." *Information Systems Research* (3:1), 1992, 60-95.
- Delone, W. H. and E. R. Mclean. *Information systems success revisited*. HICSS Conference, Hawaii, 2002.

- Galbraith, J. *Designing complex organizations*. Reading, MA, Addison-Wesley, 1973.
- Goodhue, D. L., M. D. Wybo, et al. "The impact of data integration on the costs and benefits of information systems." *MIS Quarterly* (16:3), 1992, 293-311.
- Griffin, J. "Methodologies for the future." *Software Magazine* (17:2), 1997, 5-7.
- Howard, G. S. B., T.; Janicki, T.; Liegle, J.; Klein, S.; Albert, P.; Cannon, D. "The efficacy of matching information systems development methodologies with application characteristics - an empirical study." *Journal of Systems and Software* (45), 1999, 177-195.
- Joshi, K. and M. Curtis. "Issues in building a successful data warehouse." *Information strategy: the executive's journal*, Winter 1999, 28-35.
- March, J. G. and H. A. Simon. *Organizations*, Wiley, 1958.
- Markus, M. L. "Power, politics and MIS implementation." *Communications of the ACM* (26:6), 1983, 430-444.
- Pfeffer, J. *Power in organizations*. Marshfield, Massachusetts, Pitman Publishing, 1981.
- Ribbers, P. M. A., R. R. Peterson, et al. *Designing information technology governance processes: diagnosing contemporary practices and competing theories*. HICSS conference, Hawaii., 2002.
- Sambamurthy, V., S. Venkatraman, et al. "The design of information technology planning systems for varying organizational contexts." *European Journal of Information Systems* (2:1), 1993, 23-35.
- Thompson, J. D. *Organizations in action*. New York, NY, McGraw-Hill, 1967.
- Tushman, M. L. and D. A. Nadler. "Information processing as an integrating concept in organizational design." *Academy of Management Review* (3:3), 1978, 613-624.
- van Slooten, K. and B. Schoonhoven. "Contingent information systems development." *Journal of systems software* (33), 1996, 153-161.