

December 2002

CURRENT ISSUES IN DATA WAREHOUSING

Keith Lindsey
University of Memphis

Mark Frolick
University of Memphis

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

Recommended Citation

Lindsey, Keith and Frolick, Mark, "CURRENT ISSUES IN DATA WAREHOUSING" (2002). *AMCIS 2002 Proceedings*. 7.
<http://aisel.aisnet.org/amcis2002/7>

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.

CURRENT ISSUES IN DATA WAREHOUSING

Keith Lindsey
The University of Memphis
kllindsay@memphis.edu

Mark N. Frolick
University of Memphis
mark@frolick.net

Abstract

This study examined recently published cases of data warehouse failure in an attempt to find a set of factors that are commonly associated with data warehouse project failure. In addition, the purported rate of data warehouse failure is questioned, a possible cause for the high reported rate of failure is offered, and a new definition of data warehouse failure is provided.

Keywords: Data warehouse failure

Introduction

There is a common perception that the failure rate of data warehousing projects is 70 to 80 percent (Inmon 2001) and one study reported a 90 percent failure rate (Conning 2000). Given the high failure rate, it was expected that there would be a significant number of data warehouse failure cases reported in the academic and practitioner literature. This study was undertaken to perform a detailed analysis of these cases to determine whether the presence (or absence) of any specific factors or combination of factors might be correlated with instances of failure.

Watson et al. (1999) reported the details of eight data warehouse projects that were considered failures. Beyond those, only nine additional cases of data warehouse failure have been reported since 1995 and many of these did not contain information that described the mode of failure. The simultaneous lack of documented failures juxtaposed with the common perception that most data warehouse projects are failures does not stand to reason.

Common Failure Factors

Though the evidence was anecdotal, the new cases of data warehouse failure were evaluated and the reasons for failure were found to fit within the framework provided by Watson et al. (1999). Common failure factors included weak management and inadequate user involvement. Scope creep and turnover of organizational personnel did not appear to be as common or as detrimental as previously considered.

Ramifications of the Reported Failure Rate

About 3,000 data warehousing projects are undertaken each year. If the lowest perceived data warehouse failure rate (70%) is accurate, then each year there are 2,100 failures. Yet, with the inclusion of the cases analyzed by Watson et al. (1999), only 17 cases have been reported since 1995. The discrepancy calls into question the credibility of the industry statistic.

In addition, the reported rate of failure continues to rise, indicating that data warehouse developers are unable to capitalize on improved hardware, software, and system development tools and methods. Finally, it is inconceivable that businesses continue to suffer losses amounting to billions of dollars each year without widespread analysis of these losses.

Definition of Failure

A possible reason for the discrepancy between the high documented rate of data warehouse failure and low documented instances of data warehouse failure is the ambiguity of the phrase “data warehouse failure”. A standard definition used for surveys and in industry periodicals could help bridge this gap. Our proposed definition is “the formal or informal cancellation of a data warehouse project, or the condition where a completed project’s performance falls significantly short of the business need.”

Conclusions

A review of industry statistics might cause one to expect to find details concerning hundreds of cases of data warehouse failure each year. The remarkable fact that fewer than 20 cases have been reported in the literature since 1995 precludes a rigorous analysis of the reasons for data warehouse failure, and further, it causes one to question the common perception that more than 70% of data warehouse projects are failures. A potential source of this discrepancy is the broad definition of the phrase “data warehouse failure.” Finally, it is proposed that a standard definition, based on failure to fulfill the business need, would help reconcile the simultaneous existence of two mutually exclusive industry conditions.

References

- Conning, *2000 Data Warehousing and Data Mining in the Insurance Industry: Floods of Information, Fountains of Knowledge*, Conning & Company Insurance Research and Publications, 2000.
- Inmon, W. “Information Management: Charting the Course: Little White Lies,” *DM Review*, August 2001, online at <http://www.dmreview.com/master.cfm?NavID=198&EdID=3819> accessed 8/21/01.
- Watson, H., Gerard, J., Gonzalez, L., Haywood, M., & Fenton, D. "Data Warehousing Failures: Case Studies and Findings," *Journal of Data Warehousing*, (4:1), 1999, pp. 44-55.