# Managing a Data Warehouse Environment

**Lecturer Corina Raduescu**
**Business School, University of Queensland,**
**Brisbane, Australia**
**c.raduescu@business.uq.edu.au**

REZUMAT. În mediul global competitiv de faceri din zilele noastre, este crucial pentru organizații să înțeleagă și să gestioneze toate informațiile din interior pentru a lua decizii în timp real și a răspunde la schimbările continue din mediul de faceri. Ca o consecință, multe organizații se orienteaza spre o folosire exhaustivă a datelor pe care le dețin, prin folosirea conceptului de „depozit de date" (Data Warehouse, DW).

## 1 Introduction

Since the early 1990s, DWs have become the technology of choice for building data management infrastructures. As a result, DW is near the top of most companies' strategic IT initiatives[1] and has emerged as one of the most powerful tools in delivering information for users, providing a powerful advantage in today's competitive market-space. The need of DW has been felt like never before, and the market survey for DW solutions points out that the worldwide data warehousing market will expand at a compounded annual growth rate of 43 per cent, reaching $148 billion by 2003[2].

A DW consumes a significant share of corporate IT budgets and can contribute significantly to meeting information management goals. Given the importance of a DW, deploying a DW solution is not a typical IT decision therefore managerial issues related to DW are increasingly addressed. The IT leadership is critical for infrastructure and architecture

[1] http://www.tekptnr.com/tpi/tdwi/review/bboar1.htm - "Understanding Data warehousing Strategically", written by B. Boar; Distributed by The Data Warehousing Institute.
3 http://www.ciol.com/content/e_ent/data_ware - Enterprise Special - Data warehousing.

decisions, thus effective IT governance helps enterprises deal with complexity, such in a case of a DW environment. This governance has a real impact on business goals and involves top-level collaborative decisions between business and IT executives. Therefore CIO will be the most suitable person to fulfill this IT leadership role within a company. The purpose of this report seeks to examine the managerial issues from a CIO point of view in managing a data warehouse environment. The major contribution of CIO in addressing these issues, could lead to improvement of DW projects success.

A data warehouse is more than an archive for corporate data and more than a new way of accessing corporate information. A DW is a subject-oriented, integrated, time variant and non-volatile collection of data in support of management's decision making[3]. A data warehouse is designed so that its users can recognize the information they want and access that information using simple tools. The key today is in the last part of the definition: data warehouse must support enterprise decision-makers.

The purpose of a DW is to establish a data repository that makes operational data accessible for analytical processing activities such as decision support, EIS, MIS, ESS, GDSS and other end-user applications. Accordingly, DW posses some important characteristics:

- Provide consistent and reliable data
- Operational data are aggregated into summaries, when needed
- Support time-series and trend analysis
- Use a relational structure
- Use client/server architecture
- Are not normalized – data can be redundant
- Metadata are included
- Sources can be external or operational data

A DW is typically a blending of technologies, including relational and multidimensional databases, client/server architecture, extraction/transformation programs, graphical user interfaces, and more (see Figure 1 Appendix for architecture). A well-defined and properly implemented DW can be a valuable competitive tool. DW has its own unique peculiarities and characteristics that make its development unlike the development of other applications. Not every enterprise is able to

[3] http://www.cait.wustl.edu/cait/papers/prism/vol1_no1 - "What is a DW", written by W.H. Inmon; Distributed by Prism Solutions, Inc.

successfully develop an effective DW, in fact there are many more failures than successes[4].

## 2 Data Warehouse Benefits

Being an intelligent enterprise is not an option: is a requirement[5]. In order to support the benefits that a DW environment could bring into organizations few case studies are presented in Appendix 1. In summary the most important benefits are:

- Business analysts can dig into the data, using data access and software including OLAP tools, statistical modeling tools, geographic information systems, data mining and data visualization tools.
- Rapid front-line decisions – ensuring end-user empowerment.
- Consolidated view of corporate data - data integration and ability to offer a "single version of the truth".
- Improved cost efficiency.
- Improved data quality and security of information.
- Increased competitive position (market share), ROI or IRR and customer satisfaction.

In order to obtain these benefits it is very important to achieve the goals of a DW. Industry watchers have noted that the success rate of DW implementations has not been exemplary[6]. However, there are things that executives can do to improve their chance of success. DWs, like all information systems, face technical and organizational challenges that affect the success or failure of their implementations. In the following part I will discuss the role of critical implementation factors (CIF) of a DW and CIO's agenda related to them.



[4] http://www.ciol.com/content/e_ent/data_ware - Enterprise Special - Data warehousing.
[5] http://www.ciol.com/content/e_ent/data_ware - Enterprise Special - Datawarehousing projects are business critical.
[6] Mukherjee, D., and D'Souza, D. "Think Phased Implementation for Successful Data Warehousing," *Information Systems Management* Spring 2003, pp. 82-90.

## 3 Critical Implementation Factors

DW is not essentially about technology but about the application of technology to business. Successful implementation of a DW requires, among other things, a significant investment of time and energy of many of those who will be its users to ensure the end results meet their needs. Thus, integration with the business strategy should be emphasized more than technological components. One way to improve success is to avoid factors that adversely affect DW implementations. When one compares lessons learned from failed and successful implementations, it becomes evident that both have a lot in common. Based on this, I put together the factors and compiled a shortlist of what I call *"Critical Implementation Factors"* (See Appendix 2 for a description of factors in each implementation phase). I grouped these factors in two main categories: technical and organizational, and discuss CIO involvement related to these factors.

### 3.1 Organizational Factors

- CIO primarily concern is to be convinced that there are real *business needs* for developing a DW – CIO must be aware of the strategic focus of a DW – goals and objectives must be ranked to help in prioritizing them and analyze potential future improvements of the system. Throughout the development these goals must be continuously reassessed.
- As a strategic investment, the *IT alignment with business*, IT investment and prioritization are areas where critical top-level decisions need to be made.
- *Executive sponsorship* is characterized by CIO ability to envision the DW rollout and to build trust and consensus. Sponsorship is the catalyst that enhances *communication* and provides support and resources for implementation – manage multiple users with different needs. The sponsorship must make available training and other resources for the continuously growth and evolution of the system.
- CIO must encourage *end user involvement and support* during the development and implementation to prevent a lack of understanding users' needs and a mismatch between business goals and technology provided.

- Good management of *organization culture*, *expectations, politics and resistance to change*[7] will lead to a smooth IT transition, system acceptance and a new vision of the IT within organization – these factors must be addressed appropriately as the project transitions across phases.

### 3.2 Technical factors

- An *audit of the technology available, IT personnel expertise and user proficiency* in dealing with technology and architecture required for such an environment – CIO may decide to use external consultants to help filling the gaps identified in this audit or to collaborate with other similar organizations.
- CIO is mainly responsible to manage the project for overall *adequate performance* of the system – ensure the quick delivery of benefits to the compartments.
- *Data access* from legacy systems must be addressed by CIO through a transition plan of conversion and software re-engineering of old data files.
- *Data quality and integrity* become strategic issues[8]. Data administrators, developers and IT managers have more options in using data profiling automation and focus on building contingency plans.
- *Disaster recovery* must be established through a consistent recovery plan to ensure the continuity of the business processes in case of disaster – IT managers must address these issues as part of risk management approach.
- *Security* of strategic and confidential data, information, and knowledge – different and restricted levels of access to corporate data for different categories of users - unethical and legal issues and their effects on users and competitors. Take into account the impact of new technologies, such as wireless LANs.
- A portfolio approach would be appropriate for a DW investment decision: *in-house or outsourcing development* – most of company

[7] CIOs need to immerse themselves in the management of that change, easing themselves, their staff, their users and their managers through the process and ultimately delivering the benefits.

[8] A recent PricewaterhouseCoopers study showed that corrupt data costs global businesses over US$ 1.4 billion a year in lost revenues, overpayments, mismanaged inventory, duplicated purchases, lawsuits and lost customers.

executives should collaborate together in identifying all potential benefits especially intangible for a better justification for time and money efforts.

- *Integration issues of business intelligence and DW* – they should be integrated with business processes to ensure the capabilities and comparative data for decision-making process.

### 3.3 Major Risks

- *Infrastructure and operations* – because the DW is primarily about building an infrastructure that the enterprise can leverage to build robust decision-support applications.
- *Applications* – breaking projects into smaller, more manageable chunks reduces risks of failures – build data marts for different departments and integrate them finally into DW.
- *Development methodology* – an incremental phased approach (prototyping) is required when user involvement is highly used.
- *Underestimation of development and maintenance costs* is a very high risk. Therefore CIO together with other top executives must develop an accurate cost-benefit analysis and understand that some benefits can be less than expected or may not be recorded immediately after implementation.

## 4 DW Current Status and Perspectives

In addition to previous critical factors and major risks, I identified few other issues that must be addressed in DW strategies.

- *Storage Issues* - companies face the need to double their storage capacity[9] every six months and they also need to interconnect isolated pools of storage in order to maximize ROI from expensive ERP, CRM and e-business systems. It seems that IS departments will become storage service providers. That will bring CIO in the center of action, by monitoring the monetary value to storage and identify solutions (See Appendix 3) - the big goal will be to formulate and automate storage policies.

[9] The prediction is that storage level will be a 5 million terabyte elephant by 2005. US CIO's exclusive survey "Managing Storage" found that on average 22% of a company's total budget will be allocated to storage.

- *Mission Critical DW* - will be the trend in the future for a CIO agenda. Initially DW were justified for decision making support, but it has found that the highest payoff comes when DW are integrated with operational systems, business intelligence and SCM – by creating Smart Supply Chains, to improve the performance of staff, systems and enterprise as a whole
- *Enterprise DW* – the idea is to provide business data directly to the business users who are best able to use it for business intelligence, business knowledge and business success. The key success of EDW lies in its implementation – drafting a blue print of what the warehouse should be and taking step-by-step approach to that vision
- *Proactive DW* – next generation systems will not wait for analysts and users to ask questions, they will be proactive, asking questions themselves. They will work like "detect and alert" systems.
- *Object-oriented DW* – the support of non-traditional data types in the same DW as traditional data types is a perfect application of object technology.
- *Dynamic Query Optimization* – a significant downside to DW are queries that never return or are cancelled. Server defined resource control capabilities can also prevent runaway queries.

## 5 Conclusions and Recommendations

DW is primarily the domain of large organizations, being considered an important infrastructure, not a single and isolated project. It is the most powerful tool in delivering information for users, by providing competitive advantage in today's market-space and restructuring IT strategy.

DW can improve business knowledge, enhance customer service and satisfaction, facilitate decision making, and help in streamlining business processes, but over 50% of DW projects failed to achieve their goals[10]. Cost of implementation and maintenance are very high and a DW is still a high-risk proposal, with new capabilities for enterprise, new responsibilities for CIOs and, of course, lots of new problems. However, with al difficulties in developing and managing a DW environment, DW market as predicted would continue to grow.

[10] Failures: DW does not meet the business objectives or ignore the users' expectations, and continuous addition of features to the system or changing scope of projects progressively over time.

Most of recommendations were addressed in discussing critical implementation factors. In addition few more recommendations will prove to be useful to be taken into considerations by IT managers.

- DW managers must consider seriously critical implementation and success factors (both organizational and technical) – do not build DW from scratch.
- Recognition and use of the incremental development approach (prototyping as a phased implementation method) together with active end-user involvement.
- Reflections on the consequences of not aligning DW to organizational structures and processes.
- During development think forward: What is the next step in DW? – A Proactive Enterprise DW.

A DW must be dynamic, flexible and extensible in order to meet the changing requirements of users and business needs as a whole. DW managers need to explore avenues that provide them with the opportunity to manage risks, improve operational efficiency, and advance the probability of success of DW projects.

**References**

[BM03] **Benko, C., and McFarlan, W.F.** *Reducing Risks by managing Chunks*, http://www.cio.com.au/ 2003.

[Bro03] **Broadbent, M.** *A Clear and Present Objective*, http://www.cio.com.au/ , 2003.

[Cio02] **CIO, Staff.** *What Elephant?*, http://www.cio.com.au/ 2002.

[Con01] **Connor, D.** *Report: DW failures Commonplace*, http://www.nwfusion.com , 2001.

[Con02] **Connx Solutions**, *Connx Case Study: Texas A&M University* http://www.connx.com/products/casestudy/pdf/Texas-AM.pdf

[CJ98] **Curtis, M.B., and Joshi, K.** *Lessons Learned from the Implementation of a Data Warehouse*, Journal of DW (3:2) 1998, pp. 12-18.

[Hac03] **Hackathorn, R.** *Current Practices in Active Data Warehousing*, http://www.bolder.com/case_studies/wid489.pdf , 2002, pp. 24.

[Hap02] **Hapgood, F.** *Big Just Got Bigger*, http://www.cio.com.au/, 2002.

[Hea02] **Head, B.** *Altered States*, http://www.cio.com.au/, 2002.

[Hit02] **Hitchman, S.** *Clean Up Your Act*, http://www.cio.com.au/, 2002.

[HL02] **Hsieh, C. T., and Lin, B.** *Web-based DW: Current Status and Perspective*, Winter 2002-2003, pp. 1-8.

[Inf03] **Informatica, Inc.** *Success Stories: "George Washington University DW*, ***http://www.informatica.com/customers/customer+success/gwu_07_02_cs1333_a.pdf***

[JT00] **James, A., and Thompson, S.H.T.** *Management in DW: Insights from The Housing Development Board*, Decision Support Systems (29) 2000, pp. 11-20.

[Kel97] **Kelly, S.** *W in Action*, John Wiley & Sons 1997.

[MD03] **Mukherjee, D., and D'Souza, D.** *Think Phased Implementation for Successful Data Warehousing*, Information Systems Management Spring 2003, pp. 82-90.

[Sin02] **Sinn, B.** *Enterprise DW Is the Next Step in DW*, November, 2002, pp. 11-12.

[WGH99] **Watson, H.J., Gerard, J.G., Gonzales, L.E., Haywood, M.E., and Fenton, D.** *Data Warehousing Failures: Case Studies and Findings*, Journal of DW (4:1) 1999, pp.44-55.

[WWW03]**Enterprise Special** *Datawarehousing projects are business critical*, http://www.ciol.com/content/e_ent/data_ware

## Appendix 1

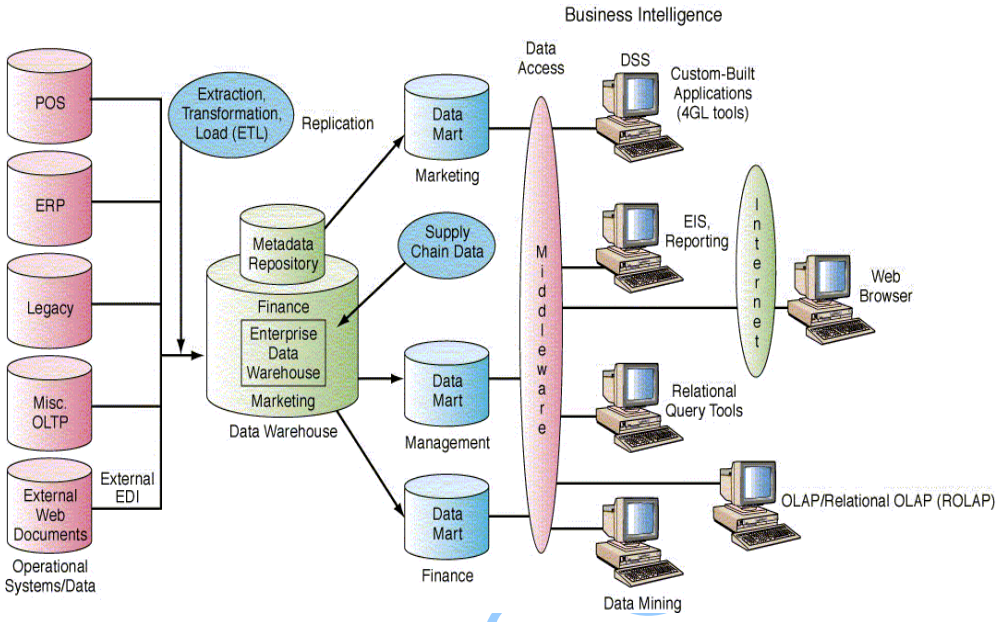


*Figure 1 – Data Warehouse Architecture*

## Case Studies

### 1. The Internal Revenue Service

– analysis for US Treasury: uses Compliance DW

Business issues:

- Billions in uncollected revenue
- Improve decision-making process, research
- Increase productivity
- Identify "upstream" procedures to improve service to tax payers

Technical issues:

- How to manage terabytes of data
- Inability to access corporate data or perform complex analysis

Business Results*:*

- ROI of 250 to 1
- Analyze trends and previously unknown patterns

- Determine rapid-response revenue protection strategies through detection of fraud
- Increased productivity by producing results more quickly at lower costs

**2. American Airlines**
– ticket sales data system Teradata
Business issues:
- Maximize revenue
- Enable business analysts to identify opportunities for competitive advantage

Technical issues:
- Increased data storage capacity
- Improve openness, scalability, and performance
- Reduce IS costs

Business Results:
- Fraud detection recovered $60,000 the first week the system was running
- Identifies when fees are not paid
- Quick and easy analysis of ticket sales
- Save $200,00 per year in IT costs
- Significantly increased query response

**3. Bank of Montreal**
– new customer relationship management system
Business issues:
- Identify and retain the most profitable portion of its customer base
- Attract new customers with similar profiles
- Increase overall effectiveness of marketing programs and overall profitability

Technical issues:
- Access to multiple information systems – data integration

Business Results:
- IRR of 100+ percent
- Payback period of less than 2 years
- Market share up by more than 60%
- System up in four months and quick query response

## Appendix 2

| | CIFs | Primary Emphasis during . . . Pre-Implementation Phase | Implementation Phase | Post-Implementation Phase |
|---|---|---|---|---|
| 1. Technical | Data | Conducting a data audit | Establishing processes and policies | Setting standards |
| | Technology | Conducting a technology audit | Designing architecture | Setting technology standards |
| | Expertise | Conducting a workforce audit | Developing a hiring plan | Setting workforce standards |
| 2. Management | Executive sponsorship | Envisioning | Communicating | Supporting |
| | Operating sponsorship | Supporting | Coordinating | Sustaining |
| 3. Goals and objectives | Having a business need | Prioritizing | Delegating | Reassessing |
| | Clear link to business objectives | Focusing | Identifying | Measuring |
| 4. User | User involvement | Consensus building | Participating | Evolving |
| | User support | Approving | Participating | Reassessing |
| | User expectation | Building | Enhancing | Altering |
| 5. Organization | Organizational resistance | Managing change anxiety | Managing turmoil | Managing adjustment anxiety |
| | Organizational politics | Resilience | Resilience | Resilience |
| 6. System | Evolution and growth | Clarify goals and scope | Flexible response | Support early adopters |

## Appendix 3

For CIOs who are just beginning to evaluate storage with an eye to drawing up a unified plan for the future, some tantalising new developments are on the horizon. However, they come with a couple of caveats. Not all the tools are in place yet. And standards squabbles between hardware vendors are stalling important interconnectivity projects.

*1. Storage resource management (SRM)* tools, currently available in early rollouts, keep a bird's-eye view of storage capacity on the network and allocate more capacity as needed. If the SRM software sees that more storage is needed by the CRM applications, it will locate and reallocate some unused capacity.

*2. Storage network management (SNM)* tools create a map of all the devices on the storage network and monitor for errors, such as network or server failure, automating a manual process.

*3. Storage virtualisation software* acts as the Adobe Postscript for storage - it fools a variety of proprietary devices into thinking alike, thus vastly increasing interoperability among storage devices.