

Internal control risks within the data warehouse environment

SP de la Rosa

School of Accountancy
University of Pretoria

H de Jager

School of Accountancy
University of Pretoria

Abstract

Decision support systems enable management to improve both their strategic and short term decision making processes. An example of a decision support system is the data warehouse. A study of the sufficiency of audit resources relating to the data warehouse was conducted. It is evident that the internal audit profession has been slow in identifying the specific internal control risks relating to such an environment. With this in mind, we undertook a study to determine the nature of the specific internal control risks within the data warehouse environment. The study also provides a brief insight into what future trends and developments can be expected within the data warehouse environment. The effect that such enhancements could have on the internal auditor's assessment of internal control risk were also noted.

Keywords:

Data warehouse, dependent data mart, internal control risk, COBIT, decision support systems, information systems, distributed, heuristic, subject orientated, time variant, system development life cycle, strategic, interface, database, breadbox analysis, data model, disaster recovery, monitoring, internal audit

1 Background

In today's highly competitive business markets of survival of the fittest, business management seek tools and new strategies which allow them to stay one step ahead of their competitors. Decision support systems are just another management tool for attaining this objective (Weber 1982:117). This evolving technology allows for the cultivation of information-sharing within organisations thereby enabling employees to solve dynamic problems and reduce costs (Warigon 1998:55).

In a survey conducted by a South African computer magazine (Du Plessis 1998:1), data warehousing was identified as one of the top seventeen application and system areas acknowledged by information technology specialists as crucial to the survival of their organisations. In addition, another computer

article (Anon 1998:22) assessed the data warehouse market as worth \$1.47 billion in 1996 with an estimated growth to 1998 of 28%.

Internal auditors seek to align themselves with the organisation's major objectives and focus on adding value to the business process. They therefore have a unique responsibility to focus on what management considers critical to the organisation's survival and success (Ridley 1996:24). If management considers the data warehouse environment a vital core of operations, internal auditors own the responsibility of ensuring that they are able to advise management concerning the internal control risks within such an environment.

2 Purpose and reason for study

2.1 Purpose of study

The purpose of this study is to identify the most pertinent internal control risks within the data warehouse environment. Finally, the impact of future developments within the data warehouse environment on the assessment of internal control risks will be considered.

The research methodology applied in this study consisted mainly of an understanding of literature in the fields of both internal auditing and systems and applications development.

Also, to identify the most pressing issues within the data warehouse environment, an empirical study was also conducted.

2.2 Reason for undertaking the study

An extensive evaluation of audit resource materials revealed that the internal audit profession has given scant attention to the data warehouse environment. There has also been a lack of focus on what impact this evolving technology will have on the assessment of internal controls¹. This study will address internal control risks at the following stages:

- Development phase of the data warehouse.
- The established data warehouse environment.
- Dependent data mart.
- Distributed data warehouse environment.
- Future developments within the data warehouse environment.

3 Defining the data warehouse environment and internal control risk

The following section shows how the data warehouse forms part of the decision support environment. It also describes the fundamental principles of the data warehouse. As an introduction to the remainder of the study, this section will conclude by identifying the overall

¹ Assumption based on the writer's personal evaluation of internal and external audit source materials and the results of extensive world wide web searches relating to the audit of the data warehouse environment.

internal control risks that can be expected within the data warehouse environment.

Since internal control risk is the central theme of this study, we will first define it.

3.1 Defining internal control risk

Internal control risk is the risk that management's plans, organisation and associated procedures will not provide reasonable assurance that the organisation's goals and objectives will be achieved (IIA 1983). We will also rely on COBIT's (ISACA 1998) information criteria identified by the Information Systems Audit and Control Association as a method of categorising risk exposure to the organisation. These criteria are:

- *Effectiveness*: Information being relevant and pertinent to the business process as well as being delivered in a timely, correct, consistent and usable manner.
- *Efficiency*: Provision of information through the optimal (most productive and economical) use of resources.
- *Confidentiality*: Protection of sensitive information from unauthorised disclosure.
- *Integrity*: Accuracy, completeness and validity of information in accordance with business values and expectations.
- *Availability*: Information being available when required by the business process now and in the future. It also concerns the safeguarding of necessary resources and associated capabilities.
- *Compliance*: Complying with those laws, regulations and contractual arrangements to which the business process is subject, i.e. externally imposed business criteria.
- *Reliability*: Provision of appropriate information for management to operate the entity and for management to exercise its financial and compliance reporting responsibilities.

3.2 The data warehouse and its concepts

A data warehouse is a structure for organising information systems. This technology grew from repeated attempts by various researchers and organisations to provide businesses with flexible, effective and efficient means of

obtaining sets of data. These sets have come to represent one of the organisation's most crucial and valuable assets (Gupta 1997:15). W.H. Inmon, acknowledged as the father of the data warehouse concept, defined a data warehouse as a "subject-oriented, integrated, time variant, non-volatile collection of data in support of management's decision making process" (Inmon 1996:33).

The following key concepts are raised in light of Inmon's definition of a data warehouse (ibid.):

- *The data warehouse is subject orientated*
Data stored within a data warehouse is physically separated from its source. The source which is operational in nature could include data relating to anything such as aviation records, hamburger sales or even telephone calls placed by subscribers.
- *The data warehouse has integrated components*
All data stored must be integrated to ensure the success of the data warehouse. Naming conventions, measurements of variables and encoding structures must be consistent and no variations should be able to be entered into the data warehouse's repository.
- *The data warehouse is time variant*
Data retained within a data warehouse covers a long time span - anywhere from five to ten years. Conversely in traditional applications, current values cover nothing more than sixty to ninety days (this is considered sound design for applications). Every group of data retained within the data warehouse includes some element of time, such as day, week, month, etc. This time element may even be implicit in nature.
- *The data warehouse is non-volatile*
Once the initial loading of data is complete, the only other operation performed within the data warehouse environment is access to the data. In stark comparison, inserts, deletes and amendments to data within the traditional application are normal and usually occur on a record-by-record basis. This simplification within the data warehouse expedites the development process in that control over changes to data is no factor.

3.3 *General internal control risks within the data warehouse environment*

According to the COBIT control criteria identified above, the major internal control risks which could be faced by an organisation employing a data warehouse, could have material effects on the successful operation of the organisation. Some indication of the major internal control risks faced by an organisation embarking on an implementation, or who might currently possess such an environment follows (Curtis & Joshi 1997:40-44):

- *Effectiveness*
Effectiveness is arguably the most important criteria for a data warehouse. Not only must the information be pertinent to the user, but timeous and consistent delivery of data is imperative. Since data retained within the warehouse is obtained from a multitude of sources, untimely and inconsistent delivery of data could skew results relied upon by management. Resultant misinformed decisions could be responsible for the demise of an organisation.
- *Integrity*
The integrity of data relates to the accuracy and completeness of information. This could involve failing to ensure that data received from source systems is cleansed before being permitted access to the data warehouse environment. If information retrieved from the data warehouse proves to be incorrect, strategic management decisions derived from such inaccurate data could also negatively impact the organisation's operations.
- *Availability*
Since data warehouses usually retain unbelievably high volumes of data which are combined to present consolidated results, ineffective storage could result in management not being able to obtain data which can be effectively analysed. A similar scenario is true if one considers that availability is also impacted by ineffective backup or incomprehensive disaster recovery procedures.
- *Efficiency*
As management seek to minimise costs and achieve optimal return on investment, warehouses can become uneconomical over time if not effectively monitored for efficiency.

Inefficient data warehouses storing superfluous data can adversely impact the ability of the database resources to make information available in the form and time frame needed by the user.

- **Confidentiality**
The data warehouse is a clearly categorised repository of data, usually storing information by degree of importance. Intruders therefore find this environment irresistible. Leaking of past patents, company strategies and other sensitive information could leave an organisation doomed. Physical security risks may also negatively impact the data warehouse environment.
- **Reliability of information**
Ineffective planning of the data warehouse could leave management with a model which does not supply the information required for improved management decision making. Not only would this result in a waste of valuable company resources, but will also deter future management from ever considering data warehousing as a tool for strategic thinking.

data warehouse is provided to the end-user. It is vital that appropriate planning and preparation take place before embarking on such an extensive exercise. If appropriately planned, the project will result in significant returns and a more controlled data warehouse environment (Kachur 1999:4).

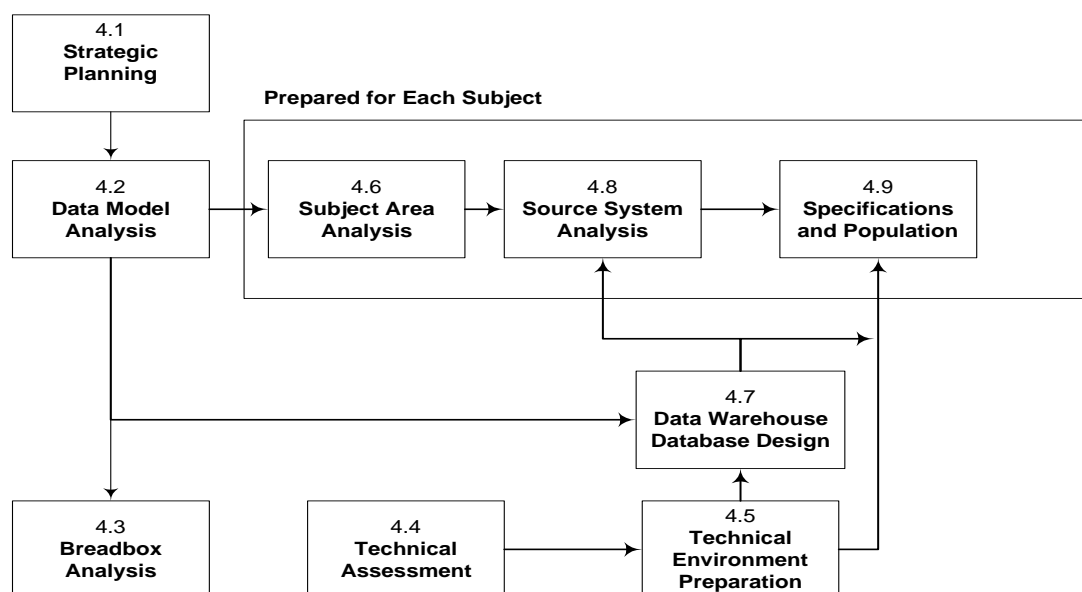
4.1 *Why a different system development life cycle exists for data warehouses*

Data warehouse development is heuristic (Inmon 1996:73). This means that the development and criteria of the subsequent phases of such a project are dependent on the outcome/results of previous phases within the development cycle. The internal auditor must realise that this is why the traditional system development life cycle cannot be applied to data warehouse development: The exact usage requirements for the data warehouse will not be known until the data warehouse environment has been populated with data. Although management and the development team may estimate what usage they expect to derive from the data warehouse environment, they must avoid making detailed assessments until populated data has been made available (ibid.).

4 Data Warehouse Development

Effective project management is a key ingredient for ensuring that a well controlled

Figure 1 - System development life cycle for the data warehouse



Source: Inmon, 1996: 350

4.2 Interface development

Inmon's data warehouse development life cycle is reflected in Figure 1. The study relies on Inmon's framework as a means of identifying internal control risks. A brief description of the various key internal control risks within each of the phases is detailed below:

4.2.1 Strategic planning

According to COBIT's information criteria identified in chapter 1, risks within the above mentioned process affect the reliability criteria of information.

The following detailed internal control risks are identified:

- By not adopting a project framework specific to the data warehouse development, the efficient and effective implementation of the data warehouse will be hampered.
- Resources will be wasted if an initial pilot project is not run to ascertain whether data warehouse benefits will be realised on a smaller scale.
- Incomplete or inaccurate project justifications could result in the organisation not realising the expected benefits of the data warehouse development.
- An inaccurate cost administration framework could result in expected project benefits not being clearly identified and accounted for.
- The appointment of an incomplete project team could result in key processes during the data warehouse development not being addressed (this includes the establishment of the data administrator role within the organisation).

4.2.2 Data model analysis

According to COBIT's information criteria identified in chapter 1, the risk within the above mentioned process affects the integrity and reliability criteria of information.

The following detailed internal control risk is identified:

- The data model must be applied in the development of the data warehouse. If not, the project team may not detect all major subject areas and ensure

that the data warehouse relies on the most accurate data available from various source systems.

4.2.3 Breadbox analysis

According to COBIT's information criteria identified in chapter 1, the risk within the above mentioned process affects the integrity and reliability criteria of information.

The following detailed internal control risk is identified:

- An incomplete assessment of the volume of data which will be retained by the data warehouse could result in:
 - i. Insufficient or excessive hardware being purchased.
 - ii. User needs not being met (by not ensuring that a sufficient level of data detail can be provided by the data warehouse).

4.2.4 Technical assessment

According to COBIT's information criteria identified in chapter 1, the risks within the above mentioned process affect the availability and reliability criteria of information.

The following detailed internal control risks are identified:

- An incorrect technical assessment could result in the final data warehouse not being able to handle the expected data volumes.
- Information will not be delivered to the user on a timely and consistent basis.

4.2.5 Technical environment preparation

According to COBIT's information criteria identified in chapter 1, the risk within the above mentioned process affects the effectiveness and availability criteria of information.

The following detailed internal control risk is identified:

- The organisation's existing applications and Information Technology operations could be negatively impacted by the introduction of the data warehouse environment if the technical

environment is not suitably prepared for the structural configuration.

4.2.6 *Subject area analysis*

According to COBIT's information criteria identified in chapter 1, the risks within the above mentioned process affect the effectiveness and reliability criteria of information.

The following detailed internal control risks are identified:

- The user's needs will not be met if the most appropriate subject areas are not chosen, based on the details contained in the data model analysis.
- Inaccurate and untimely data reliance could occur if the most appropriate subject data is not selected in instances where subject data can be retrieved from multiple sources.

4.2.7 *Data warehouse design*

According to COBIT's information criteria identified in chapter 1, the risks within the above mentioned process affect the effectiveness, integrity and reliability criteria of information.

The following detailed internal control risks are identified:

- An incomplete data warehouse design could produce incomplete and unreliable data elements in the final data warehouse environment.
- Misinformed end user decisions could result which, depending on the significance of the data warehouse query, could directly affect the viability of the organisation.

4.2.8 *Source systems analysis*

According to COBIT's information criteria identified in chapter 1, these risks within the above mentioned process affect the integrity, effectiveness and reliability criteria of information.

The following detailed internal control risks are identified:

- An incomplete source system analysis could result in incorrect data elements being transferred to the data warehouse.

- Non-existence of a conversion plan could result in team participants being unaware of the approved standards and conversion requirements applicable to the development of the data warehouse.
- Without a source systems analysis, it may prove difficult for the project team to trace incorrect data residing within the data warehouse environment back to the operational system that created the errant data.

4.2.9 *Interface specifications and population*

According to COBIT's information criteria identified in chapter 1, the risks within the above mentioned process affect the effectiveness, reliability and integrity criteria of information.

The following detailed internal control risks are identified:

- The overall reliability of the final data elements could be affected by the incorrect transfer, loading and analysis of data prior to migration to the data warehouse environment.
- Unnecessary errors and the loss of data integrity could occur without an approved methodology for developing programmatic code for the extraction of data from source systems to the data warehouse.
- Non-existent conversion specifications could be responsible for certain data elements not being identified by the project team.

5 **Established Data Warehouse Environment**

In this section we identify six internal control risks which may exist within an established data warehouse environment. Under each of the risks identified we indicate which of COBIT's information criteria, viz. effectiveness, integrity, availability, efficiency, confidentiality are affected.

5.1 *Inability to measure data quality and ensure satisfactory refreshing of data*

Without continuously monitoring data quality, management cannot ensure that data complies with approved management standards (Bohn 1997:1).

The refreshing of data within the data warehouse is fixed during the codification of the interfaces between the source system and data warehouse applications (Inmon 1996:280). If the refreshing rate of data is not frequently revisited with the user, such rates may become unsuitable in the future and result in users placing reliance on inaccurate data presented by the data warehouse.

According to COBIT's information criteria identified in chapter 1, this risk affects the effectiveness, integrity, availability and efficiency aspects of information (Curtis & Joshi 1997:40-43).

5.2 *Not ensuring the completeness of data migrated to the data warehouse* (Fryman 1997:46)

In instances where management decides to include additional subject areas over time, ineffective project management and the lack of an approved development methodology will result in new subject areas not being included in the most efficient and effective manner.

Changes made to source systems without considering the data warehouse environment could affect the completeness of data migrated to such an environment (Inmon 1996:182). An ineffective communication process amongst the various Information Technology teams and end-users can result in these changes not being communicated effectively.

According to COBIT's information criteria identified in chapter 1, this risk affects the integrity aspect of information (Curtis & Joshi 1997:40-43).

5.3 *Ongoing availability of data warehouse operations cannot be ensured* (Warigon 1998:55)

A 1997 survey (Semer 1998:47) which tracked the major classifications of data loss among 50 000 organisations was conducted by Ontrack Data International Inc.. The survey indicated that 44% of data losses were caused by hardware or system malfunctions.

According to the Disaster Recovery Journal (Semer 1998:42), organisations suffered significant disaster-related costs in 1997:

- Each on-line outage averaged four hours and cost American companies an average of \$329 000 in lost revenues and productivity.

- For each hour of unscheduled downtime, 355 worker hours were lost.
- Major businesses lost 38.1 million work hours, or \$444 million in wages annually.

It is apparent from these statistics that a significant risk is faced by organisations should mission critical systems become unavailable. Although the data warehouse environment is only one source of user information, the statistics are an indication of potential losses. Without consistent and supported data warehouse services, the end user may be unable to make informed management decisions.

According to COBIT's information criteria identified in chapter 1, this risk affects the availability aspect of information.

5.4 *Overall data warehouse administration becomes ineffective and inefficient* (Warigon 1998:59)

If the data warehouse environment is not effectively monitored, it can become unwieldy. In many instances, ineffective and inefficient data warehouses are caused by not regularly archiving outdated data and by not executing frequent capacity planning measures. Not performing these activities ultimately leads to increased annual storage, processing and operating costs (ibid.).

Routine archiving of data involves the rolling up of outdated data to higher levels of summary (Inmon 1996:69). This rolling of data can either be by means of transferring data from one level of the data warehouse structure to another or retaining data within a high-performance storage medium.

According to COBIT's information criteria identified in chapter 1, this risk affects the efficiency, effectiveness and availability aspects of information.

5.5 *Data warehouse access is not restricted to authorised users* (Warigon 1998:55)

Unauthorised access to data retained within the data warehouse can result in significant losses to the organisation (Warigon 1998:55-60). These threats can be caused by accidental or malicious attacks from employees. Outside threats can be caused by competitors. The result of such unauthorised access could be negative publicity for the

organisation and a loss of continuity of data warehouse operations. Management will need to identify security vulnerabilities which could negatively impact the organisation's image. As part of this assessment, physical security risks should also be considered. (ibid.).

According to COBIT's information criteria identified in chapter 1, this risk affects the confidentiality aspect of information.

5.6 Ongoing risk assessments of the data warehouse environment are not conducted (Warigon,1998:55)

Cost-effective measures are required to address the most significant risks within the data warehouse environment. Organisations should be focusing on ways to limit costs and secure only mission critical assets (ibid.). This valuable information can only be obtained by performing, and revisiting, risk assessments over information technology environments such as the data warehouse. These assessments will allow management to identify how vital the risks are within the data warehouse environment and thereby apply limited resources in the most effective and efficient manner.

Ultimately, if organisations do not frequently perform risk assessments, the effective and efficient utilisation of resources cannot be ensured.

According to COBIT's information criteria identified in chapter 1, these risks affect the effectiveness, integrity, efficiency and reliability aspects of information.

6 Dependent Data Mart Environment

Although organisation wide data warehouses were the first to be introduced, they can sometimes fail to provide information to the end user effectively and within an acceptable time frame (Bersin 1996). To address these weaknesses, an organisation may choose both an organisation wide data warehouse that addresses all operations and a simplified data structure for selected business units or subjects. The key advantage of such a data structure is improved response times to access data and process queries (Bersin 1996).

In this section we identify two unique internal control risks which may exist within an established data warehouse environment.

Under each of the risks identified we also indicate which of COBIT's information criteria, viz. effectiveness, integrity, availability, efficiency, confidentiality are affected.

6.1 Insufficient periodic response time monitoring

The inconsistent or utter lack of ongoing response time monitoring within the data mart environment may result in information not being provided to the user on time. In such instances the user will either stop utilising the data mart environment or be unable to make informed management decisions (Bersin 1996).

According to COBIT's information criteria identified in chapter 1, this risk affects the efficiency, effectiveness and availability aspects of information.

6.2 Transfer of data from the organisation wide data warehouse to the data mart is not controlled

The data mart may become inefficient and ineffective if the project team fails to ensure that only the most necessary data is transferred from the organisation wide data warehouse to the data mart (Bersin 1996).

The upload of data from the data warehouse to the data mart may also create data integrity and availability problems if uncontrolled. The uncontrolled uploading of data may arise if:

- The frequency of updating data is not in agreement with the end user's needs.
- The data mart is not updated with data changes made in the data warehouse environment.
- The inability of the data mart environment to notify the users of subsequent changes made to data already relied upon, may result in incorrect management decisions being made. This is applicable in instances where comprehensive refreshing of data occurs.

In all three instances, the significant loss of decision making and the reliance on incorrect data can cause significant financial losses to the organisation.

According to COBIT's information criteria identified in chapter 1, these risks affect the

availability, integrity and reliability aspects of information.

7 Distributed Data Warehouse Environment

There are certain instances where a distributed data warehouse will be more appropriate than a centralised structure (Bell 1992:2-4):

- Experience has shown that 90% of data operations are local, meaning that in instances where organisations are dispersed geographically, the need for users to access their data locally is increased.
- For back-up purposes, it is considered good business practice to have data replicated in a number of sites to ensure continuous operation.
- Improved technology addresses limited access to data by centralised processing operations.
- Distributed data warehouses can be more easily expanded to accommodate increasing data volumes than can a corporate data warehouse (Inmon 1996:213).

We have identified three particular internal control risks which may exist within an established data warehouse environment. Under each of the risks identified we indicate which of COBIT's information criteria, viz. effectiveness, integrity, availability, efficiency, confidentiality are affected.

7.1 *Distributed data warehouse access is not restricted to authorised users*

There are two major risks concerning unrestricted access to the distributed data warehouse environment (Bell, 1992:5):

- Ensuring controlled access across open communication channels.
- Ensuring optimal access to distributed resources.

Significant risk of unauthorised disclosure of information may occur if access restrictions do not ensure that only valid users have the right to view data retained within the data warehouse environment.

An optimal access consideration involves the risk of inefficient access to data. This could result in users under utilising the data warehouse due to poor response times.

According to COBIT's information criteria identified in chapter 1, this risk affects the confidentiality, integrity and efficiency aspects of information.

7.2 *Ongoing availability of the distributed data warehouse operations cannot be ensured*

As in the case of an established data warehouse environment identified in section 3, the distributed data warehouse is also prone to expected and unexpected failure (Bell 1992:5). The most common forms of failure within the distributed data warehouse environment can be summarised as follows (Bell 1992:233-239):

- *Local transaction failures*
These failures are caused either by unforeseen transaction failures, (such as system logic errors), or by system induced failures, (such as management override of computer programs or the intentional shut-down of computer operations). The severity of these failures is usually limited since they only affect a small number of transactions.
- *Site failures*
Sites operate independently in the distributed data warehouse environment. Therefore it is possible for certain sites to be operational while others have failed (referred to as partial failures). Partial failures are considered far more hazardous than a complete failure of the distributed environment. This is because it is difficult for other sites to detect instances where other reliant sites are unavailable.
- *Media failures*
Media failures are caused by hardware corruptions. The most common of these failures occur in hard disk storage devices.
- *Network failures*
Networks are considered to be the core of efficient and effective communication between the local and global sites. Although today's networks are considered to be robust, it is possible that line failures may corrupt communications. To a large degree, the ability for system software to reroute communications has overcome this type of failure.

According to COBIT's information criteria identified in chapter 1, this risk affects the availability aspect of information.

7.3 *Efficiency of processing within the distributed data warehouse is not maximised*

The efficiency of transformation and integration of data between distributed sites is identified as one of the major risk areas in a distributed data warehouse environment (Bell 1992:5). If control mechanisms are not implemented to mitigate this risk, it is possible that users may become disillusioned when the provision of data warehouse functionality is slowed. This under utilisation of assets could result in uninformed management decisions being taken which could in turn affect the profitability and even the continued operation of the organisation.

Query optimisers can be utilised to address the controllable weaknesses relating to the efficient transfer and communication of information in a distributed data warehouse environment (Bell 1992:124). The task of the query optimiser is to govern and expedite the processing and data transmission required for responding to queries. This in turn ensures that either the total cost or the total response time for a query is minimised (ibid.).

According to COBIT's information criteria identified in chapter 1, this risk affects the effectiveness, availability, efficiency and reliability aspects of information.

8 Future Developments and Trends

The following areas have been identified as future developments which could significantly change existing data warehouse environments (DCI 1998:1-3). In addition to describing these developments, the study also identifies what effect these changes could have on internal control risks.

8.1 *Closed-loop business performance management*

- Although real-time data provision allows for timeous decision making, it can result in the user being inundated with unnecessary data. This will occur in instances where the criteria used in identifying reported trends are not established according to stringent management standards.

- If not properly managed and monitored, real-time data provision can result in the over utilisation of system resources.

8.2 *Increased access to data warehouse information*

- The risk of confidential information being disclosed to unauthorised users may be due to incomplete or inconsistent application of access rights, procedures and policies.
- Increasing the number of users who have access rights to the data warehouse environment can negatively impact the overall performance of the environment. This risk is increased if ongoing monitoring and the necessary upgrading of software and hardware is not conducted.

8.3 *Automatic removal of source data quality problems*

- The development of any procedure capable of correcting source data automatically within source applications without any form of stringent management control or audit trail is concerning. Such automatic procedures could result in unauthorised changes being made to source data.

8.4 *Re-engineering the development methodology*

- There is a risk that the development process may become so inwardly focused that system development methodologies will not consider end user needs as the primary input in the development process.

8.5 *Transferring of report and query functionality*

- The data warehouse will need to be accessed by a far larger audience. Therefore the risk of unauthorised access to privileged information is increased.
- User training and data warehouse package licenses are necessary and will increase costs.

9 Results of the empirical study

- Of the organisations who had embarked on a data warehouse development, 70% had developed a system methodology specifically for the data warehouse environment.
- The most significant contribution by internal audit departments to the data warehouse development phase related to security aspects. Only 16% of audit departments indicated that they were involved in the entire project development.
- Only 57% of internal auditors surveyed, indicated that the data warehouse development was completed on time. The most significant causes for these overruns were users requesting additional functionality and unexpected system complexity.
- Only 33% of organisations realised the full benefits of the data warehouse environment. Respondents identified unrealistic expectations as the major cause for these failures. This would seem to stem from senior management misperceptions of the functional capacities of a data warehouse environment.
- 50% of the local internal auditors surveyed indicated that their data warehouse developments were distributed in nature. This indicates that a fair number of local organisations are utilising the distributed data warehouse structure.
- 15% of the respondents indicated that they expected network security to be the area of greatest concern within the next two years. Electronic commerce, remote access and expanding networks were individually rated at 12%, while 8% of the respondents felt that the data warehouse environment would be a significant area of concern in the foreseeable future.
- Internal audit managers were requested to indicate which specific

data warehouse elements they felt would significantly impact their assessment of the data warehouse environment. 28% of the respondents indicated that the increase in users would have the greatest impact on their assessment of internal control risk. The risk of real-time information management was assessed at 24%.

10 Conclusion

The following overall recommendations arise from this study:

- The data warehouse development team should ensure that a particular system development life cycle unique to the data warehouse environment is utilised whenever such an environment is being developed.
- Internal audit should be involved throughout the development process to ensure that significant internal control risks are identified.
- Internal auditors should ensure that management have considered the effect of future developments within the data warehouse environment when assessing the overall risk of such an environment.
- Internal auditors should use the data warehouse environment to identify significant trends and irregularities when performing routine audits.

In conclusion, the data warehouse environment provides unique opportunities for the organisation to ensure a more reliable and consistent decision making process. As reliance on systems such as the data warehouse environment increases, the need for a more controlled system is increased. Management and the internal auditor should work together to ensure that all significant internal control risks are identified timeously and that measures are implemented to negate the effects of these risks.

References

- Anon 1998. Strong Warehouse Growth Ahead. *Computerweek*. 21 September 1998.
- Bell D., Grimson J. 1992. *Distributed Database Systems*. 1st Edition. Addison-Wesley Publishers Ltd. Great Britain.

- Bersin J. 1996. Data Marts and Beyond: A Pragmatic Approach to Enterprise Decision Support. International Data Warehousing Association. (<http://www.idwa.org/fall96/datamarts.htm>)
- Bohn K. 1997. Converting Data for Warehouses. DBMS and Internet Systems, Miller Freeman, Inc. May 1997. (<http://www.dbmsmag.com>)
- Curtis M., Joshi K. 1997. Internal Control Issues for Data Warehousing. *IS Audit & Control Journal*. Volume IV.
- DCI 1998. *A Data Warehousing Forecast*. 5 February 1998. DCI Inc. (<http://www.dci.com/news/1998/feb/dwtrends.htm>)
- Du Plessis M. 1998. Tough Times, Tougher Decisions. *Computerweek Supplement*. 5 October 1998.
- Gupta VR. 1997. Research paper: *An Introduction to Data Warehousing*, August 1997.
- IIA. 1983. SIAS 1 - Statements on Internal Auditing Standards: *Control Concepts and Responsibilities*. Altamonte Springs Florida USA. Institute of Internal Auditors. July 1983.
- Inmon W.H. 1996. *Building the Data Warehouse*. 2nd Edition. United States: John Wiley & Sons, Inc.
- ISACA. 1998. COBIT - *Control Objectives for Information and Related Technology*. 2nd Edition. Rolling Meadows. Illinois USA. Information Systems Audit and Control Association.
- Kachur R. 1999. Planning and Project Management for the Data Warehouse. DBMS and Internet Systems, Miller Freeman, Inc. February 1999. (<http://www.dbmsmag.com>)
- Ridley A. 1996. A Profession for the Twenty-first Century. *Internal Auditor*. October 1996.
- Semer L. 1998. Disaster Recovery Planning. *Internal Auditor*. December 1998.
- Warigon S. 1998. Data Warehouse Control and Security. *Internal Auditor*. February 1998.
- Weber R. 1982. *EDP Auditing Conceptual Foundations and Practice*. 1st Edition. Tokyo Japan: McGraw-Hill Kogakusha, Ltd.