

***DESIGN AND DEVELOPMENT OF A DATA WARE-HOUSE
FOR SCHOOL INFORMATION SYSTEM
CASE STUDY OF
HERITAGE INTERNATIONAL SCHOOL***

**BY
OCHEN MOSES PATRICK
MSCSSE/3939/62/DU**

**A graduation project presented to the School of Post Graduate Studies
Kampala International University in partial fulfillment of the
requirements for Degree Master of Science in Software Systems
Engineering**

Nov. 2008

TABLE OF CONTENTS

Cover page	i
Approval	ii
Declaration	iii
Dedication	iv
List of figures	v
List of tables	vi
Acronyms	vii

CHAPTER ONE

1.1.0 Background.....	2
1.1.1 Problem statement	4
1.1.2 General Objective	4
1.1.3 Specific Objectives	4
1.1.4 Significance of the study	4
1.1.5 Scope of the study	4

CHAPTER TWO

Literature Review	5
2.0 What is a Data Warehouse	5
2.1.1 Advantages and disadvantages to using a Data Warehouse	9
2.1.2 How to manage current and historical information within your Data Warehouse	11
2.1.3 Understanding the challenges of using Data Warehouses	11
2.1.4 Data Warehouse (DW) Database and Online Transaction Processing (OLTP)	12

2.1.5 How to protect your Data Warehouse	13
2.1.6 Objectives of a Data warehouse and Data flow	15
2. 1. 7 Modeling Multidimensionality	16
2.1.8 School Information System	17
2.1.9 Why need SIS?	19
2.2.0 More information on school Information system	19
2.1.1 Data Warehousing Integrates School Information Systems.	20
2.1.2 Data Models for Data Warehouse	20
2.1.4 The Methods for designing Data Warehouse Architecture	21
2.2.5 The Bottom-Up Approach	22
2.2.6 The Top-Down Approach	23

CHAPTER THREE

Research Methodology.....	20
3.1.0 Research Instruments	25
3.1.2 Developing a Survey	25
3.1.2 Study area	26
3.1.3 Procedure of the study	26
3.1.4 Data collection methods	26

CHAPTER FOUR

System Analysis and Design.....	27
4.1.0 Description of the Current System	27

4.1.2 Strengths of the current system	27
4.1.3 Weakness of the current system	27
4.1.4 Possible solutions to the problems of the current system	28
4.1.5 The proposed system	28
4.1.6 Requirements of the proposed system	29
4.1.7 System Design	29
4.1.8 Physical design	30
4.1.9 Database	33
4.2.0 An entity-relationship (ER) diagram	35
4.2.1 Designing the Star Schema	37

CHAPTER FIVE

System Implementation and Control ^{EVALUATION}	40
5.1.0 Program Implementation	40
5.1.2 Features of the new system	40
5.1.3 Program Testing	41
5.1.4 Sample Testing for SIS 1	42
5.1.5 Program Flow Chart	42
5.1.6 Implementation	44
5.1.7 Change over methods	44
5.1.8 User Manual	44
5.1.9 System Evaluation	45
5.2.0 Security requirements	45

CHAPTER SIX

Discussion, Conclusions and Recommendation	46
6.1.0 Top-Down vs. Bottom-Up Method	46
6.1.1 School Information System-Specific Issues	47
6.1.2 Lessons Learned	47

6.1.3 Future Recommendation.....48

6.13 Future Recommendation48

BIBLIOGRAPHY49

APPENDICES51

Appendix A: questionnaire54

Appendix B: Data Warehouse Evaluation Questionnaire54

Appendix C: script to determine the records of the table56

List of Figures

Figure 1: The 3-tier scheme Architecture	3
Figure 2- The DWQ quality framework.....	21
Figure 3: Data Flow Diagram	26
Figure 4: user login interface	31
Figure 5: Staff Login Interface	33
Figure 6: E-R Diagram for a School Database	36
Figure 7: Database Schema	39

List of Tables

Table 1: Heritage International School.....17

Table 2: Student table34

Table 3: Staff Table34

Table 4: Department34

DECLARATION

I Ochen Moses Patrick hereby declare that this special study paper is my original work. It is not a duplicate of any other scholar's work. Any materials cited here which are not my own have been duly acknowledged.

Signature *Ochen Moses*

Date *24/10/08*

Ochen Moses Patrick

APPROVAL

This project has been submitted with my approval as university supervisor

24/10/2008 J. Badae

Mr. Bada Joseph Kizito

Date

DEDICATION

I dedicate this work to my Guardians, The Mccarty family for having given me all the support during the study period. If it were not for their support, this work would not have been successful.

ACKNOWLEDEMENT

I wish to acknowledge the assistance I received from various people whose advice helped me in some ways towards the whole process.

There are many people especially all lecturers of school of post graduate studies KIU without their efforts; this research would not have become a reality. I am indebted to various individuals who read the draft copies and offered valuable criticisms and advice.

ABSTRACT

Data warehousing and business intelligence are key enablers in the drive to harness management information and improve decision-making. They underpin the recent wave of interest in data mining and customer relationship management, and create opportunities for performance improvements based on quantitative analysis at every level in an organization. Data warehousing is a relatively new field in information technology, and current research centers primarily on data warehousing in business environments. As new as the field is in these environments, only recently have educational institutions begun to embark on data warehousing projects, and little research has been done regarding the special considerations and characteristics of School Information System and the complexity of analyzing such data. Educational institutions measure success very differently from business-oriented organizations and the analyses that are meaningful in such environments pose unique and intricate problems in data warehousing. Heritage International School maintains database comprising of a number of files of school data. This is international school that has links to other international schools in the same category. There is need to design and develop school database that should accommodate past data as well as capture present data in order to build data warehouse.

According to the school administrator, the present school Data Warehouse is not sufficient, well structured to handle data, it gives poor quality information, lacks consistency, and above all it is not easy to access. This research describes the process of developing a data warehouse for Heritage International School, focusing on issues specific to School Information system and the general objective of this study is to examine the factors in the development and design of a Data Warehouse for a school Information System. Interviews, questionnaires and observations will be used to collect the required information for the Data Warehouse.

CHAPTER ONE

INTRODUCTION:

Heritage International School was established in September 1994 by five Christian missions: Finnish Free Foreign Mission, Church of God Mission, Kinderhilfswerk, Mission Aviation Fellowship and Baptist Mission of Uganda. A Memorandum of Agreement was signed by the parties. Heritage International School was to be a co-operative effort to support Christian mission personnel in Uganda. Children from both mission and non-mission families were to be welcomed.

The policy-making body of Heritage International School is the Board of Directors. This board is made up of representatives from all the sponsoring missions, as well as a non-voting parent representative from the Parents Teachers Fellowship. The Principal also participates on the Board without a vote and is authorized by the Board to provide for the day-to-day administration of the school. Heritage International School is registered under the Companies Act as a company not-for-profit, limited by guarantee and without a share capital. In addition, it is registered with and licensed by the Ministry of Education. Heritage International School is located in Kiwafu, Kampala, Uganda. The new and spacious campus includes several buildings with rooms for studies in language, math, music, science, crafts, social studies, computer, and others. The new campus also has a library and computer room, as well as playing fields and well-equipped playgrounds. The academic program at Heritage International School is designed to enable students to succeed in a variety of environments. The students come from different educational backgrounds and continue on in a variety of different options in different parts of the world. Heritage International School follows a learner centric model where each student develops their God given potential to a maximum. The founding vision of Heritage was to support Christian Mission personnel in Uganda and to provide a Christ-centered education for all children, regardless of their religious, social, or economic background. The founders also saw Heritage as an opportunity to provide not only a teaching ministry for missionaries, but also additional avenues of testimony and service outside of the school. Finally the founders of Heritage International School dreamed of a place where the influence of Christian Service would go beyond academics, impacting the entire international and inter-denominational Christian Community. We hold firmly to this vision.

Heritage International School offers quality education that equips each student to develop their God-given potential for a life of faith and service. Heritage International School plans to continue to offer high quality education for students. In addition it also desires to work toward improving its methods and practices so that it can continue to impact the lives of students for the Lord. To do this, it is the school's intention to provide the best facilities, curriculum, and faculty possible, thereby impacting the international and interdenominational Christian communities as well as our student population.

Heritage provides a top-notch education for each of its students from preschool to grade twelve. Teachers from North America, Europe, and East Africa, all professionally trained and experienced, provide quality teaching in the classroom, where students from around the world receive individualized attention due to a low teacher to student ratio of 1:10. The school, which uses primarily an American curriculum, boasts a state-of-the-art computer lab, a fine library, a spacious playground, and a football pitch. Students receive instruction in music, art, computer, French, Swahili, Physical Education, and swimming, in addition to the core curriculum subjects.(www.heritageinternationalschool.org)

1.1.0 Background

Bill Inmon, who is recognized as the “father of warehousing defines 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).

Ralph Kimball, another noted authority on data warehousing defines more simply as “a copy of transaction data specifically structured for agency and analysis” (Kimbal 1996). Both of these definitions stress that a data warehouse is a collection of data separate from the databases that support the day-to-day operating of an organization, and that is specifically designed for the purpose of producing reports. Therefore, the process of designing a data warehouse is somewhat different from the process of developing Online Transaction Processing (OLTP) database (Lester, 2003).

In Online Transaction Processing (OLTP) systems, data models describe the business operations in terms of small, predefined transactions. These models tend to be rather complex,

and the end users rarely interact directly with them, relying instead on Information Technology (IT) personnel to process reports for them

Data models have long been used as a communication tool between database developers to visualize the structure of a business or organization, but they are descriptive enough for end users to understand so that they can ensure that the model does indeed reflect their business requirements (Lester, 2003)

Figure 1 below shows the three-tier architecture of Data Processing Systems (DPS), as defined by (Hay 1997). The data models are the primary component of the conceptual level serving as the translator between the external views of the end users and the internal implementation that is the physical layer

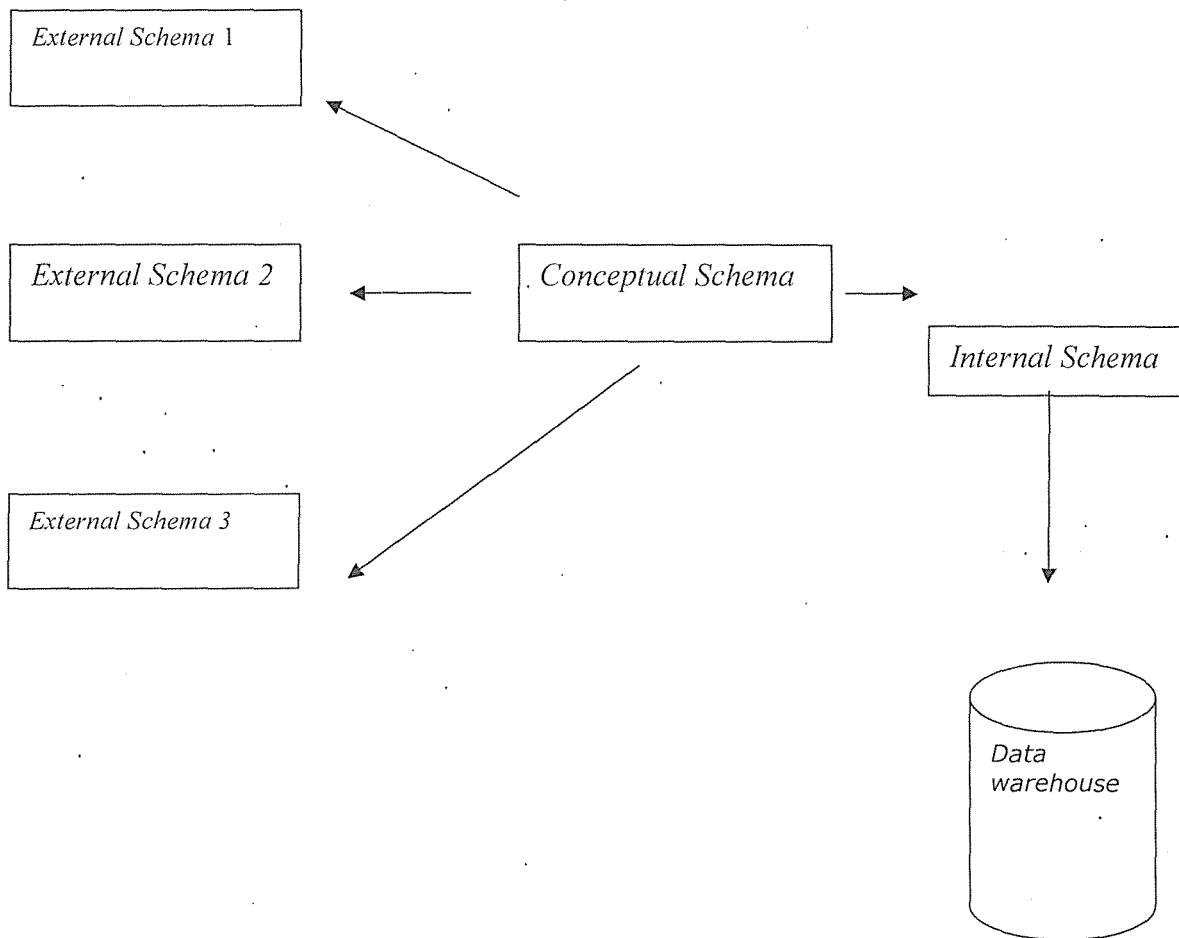


Figure 1 – The 3-tier scheme Architecture (Lester, 20

1.1.1 Problem statement:

Heritage International School maintains database comprising of a number of files of school data. This is international school that has links to other international schools in the same category. There is need to design and develop school database that should accommodate past data as well as capture present data in order to build data warehouse.

According to the school administrator, the present school Data Warehouse is not sufficient, well structured to handle data, it gives poor quality information, lacks consistency, and above all it is not easy to access.

1.1.2 General Objective

The general objective of this study is to examine the factors in the design and development of a data warehouse for School information system and to design a good data ware house that will be of benefit to Heritage International School

1.1.3 Specific Objectives of the study are:

- To investigate the Data Warehouse at Heritage International School
- To design a logical database for Heritage International School.
- To construct a database using current database technologies
- To develop the baseline measurements to evaluate the quality of the Data Warehouse.

1.1.3 Significance of the study

The creation of Data Warehouse is the basis for electronic data storage for the school. The Data Warehouse provides easy access to data whenever required by the stakeholders. There is easy data retrieval and manipulation.

1.1.4 Scope of the study:

The research will focus on a data warehousing product for Heritage International School. Data will be extracted primarily from the Students' Information System (SIS) covering a five year period.

Data Warehouse Models will be applied in an attempt to determine which one is most effective.

Finally a good Data Warehouse will be developed implemented and evaluated

CHAPTER 2

LITERATURE REVIEW

The aim of this review is to ensure that the researcher is familiar with previous work that is relevant to Data Warehouse and School Information System

Data warehouse and school information system are discussed thoroughly citing different researchers and authors.

2.1.0 What is a Data warehouse?

Ralph states that a data warehouse is "a copy of transaction data specifically structured for query and analysis". Two quibbles I have with Ralph's definition are: 1) Sometimes non-transaction data are stored in a data warehouse - though probably 95-99% of the data usually are transaction data. 2) I say "querying and reporting" rather than "query and analysis" because the main output from data warehouse systems are either tabular listings (queries) with minimal formatting or highly formatted "formal" reports. Queries and reports generated from data stored in a data warehouse may or may not be used for analysis

A Data Warehouse can be normalized or denormalized. It can be a relational database, multidimensional database, flat file, hierarchical database, object database, etc. Data warehouse data often gets changed. And data warehouses often focus on a specific activity or entity.
(www.dwinfocenter.org)

According to Greenfield (2008), the following is a list of the basic reasons why organizations implement data warehousing. This list was put together because too much of the data warehousing literature confuses "next order" benefits with these basic reasons. For example, spend a little time reading data warehouse trade material and you will read about using a Data Warehouse to "convert data into business intelligence", "make management decision making based on facts not intuition", "get closer to the customers", and the seemingly ubiquitously used phrase "gain competitive advantage". In probably 99% of the data warehousing implementations, data warehousing is only one step out of many in the long road toward the ultimate goal of accomplishing these highfalutin objectives.

The basic reasons organizations implement data warehouses are:

To perform server/disk bound tasks associated with querying and reporting on servers/disks not used by transaction processing systems

Most firms want to set up transaction processing systems so there is a high probability that transactions will be completed in what is judged to be an acceptable amount of time. Reports and queries, which can require a much greater range of limited server/disk resources than transaction processing, run on the servers/disks used by transaction processing systems can lower the probability that transactions complete in an acceptable amount of time. Or, running queries and reports, with their variable resource requirements, on the servers/disks used by transaction processing systems can make it quite complex to manage servers/disks so there is a high enough probability that acceptable response time can be achieved. Firms therefore may find that the least expensive and/or most organizationally expeditious way to obtain high probability of acceptable transaction processing response time is to implement a data warehousing architecture that uses separate servers/disks for some querying and reporting.

To use data models and/or server technologies that speed up querying and reporting and that are not appropriate for transaction processing.

There are ways of modeling data that usually speed up querying and reporting (e.g., a star schema) and may not be appropriate for transaction processing because the modeling technique will slow down and complicate transaction processing. Also, there are server technologies that may speed up query and reporting processing but may slow down transaction processing (e.g., bit-mapped indexing) and server technologies that may speed up transaction processing but slow down query and report processing (e.g., technology for transaction recovery.) - Do note that whether and by how much a modeling technique or server technology is a help or hindrance to querying/reporting and transaction processing varies across vendors' products and according to the situation in which the technique or technology is used.

To provide an environment where a relatively small amount of knowledge of the technical aspects of database technology is required to write and maintain queries and reports and/or to

provide a means to speed up the writing and maintaining of queries and reports by technical personnel

Often a Data Warehouse can be set up so that simpler queries and reports can be written by less technically knowledgeable personnel. Nevertheless, less technically knowledgeable personnel often "hit a complexity wall" and need IS help. IS, however, may also be able to more quickly write and maintain queries and reports written against Data Warehouse data. It should be noted; however, that much of the improved IS productivity probably comes from the lack of bureaucracy usually associated with establishing reports and queries in the data warehouse.

To provide a repository of "cleaned up" transaction processing systems data that can be reported against and that does not necessarily require fixing the transaction processing systems

Please read my essay on what data errors you may find when building a data warehouse for an explanation of the type of "errors" that need cleaning up. The Data Warehouse provides an opportunity to clean up the data without changing the transaction processing systems. Note, however, that some data warehousing implementations provide a means to capture corrections made to the Data Warehouse data and feed the corrections back into transaction processing systems. Sometimes it makes more sense to handle corrections this way than to apply changes directly to the transaction processing system.

To make it easier, on a regular basis, to query and report data from multiple transaction processing systems and/or from external data sources and/or from data that must be stored for query/report purposes only

For a long time firms that need reports with data from multiple systems have been writing data extracts and then running sort/merge logic to combine the extracted data and then running reports against the sort/merged data. In many cases this is a perfectly adequate strategy. However, if a company has large amounts of data that need to be sort/merged frequently, if data purged from transaction processing systems needs to be reported upon, and most importantly, if the data need to be "cleaned", data warehousing may be appropriate.

To provide a repository of transaction processing system data that contains data from a longer span of time than can efficiently be held in a transaction processing system and/or to be able to generate reports "as was" as of a previous point in time

Older data are often purged from transaction processing systems so the expected response time can be better controlled. For querying and reporting, this purged data and the current data may be stored in the data warehouse where there presumably is less of a need to control expected response time or the expected response time is at a much higher level. - As for "as was" reporting, some times it is difficult, if not impossible, to generate a report based on some characteristic at a previous point in time. For example, if you want a report of the salaries of employees at grade Level 3 as of the beginning of each month in 1997, you may not be able to do this because you only have a record of current employee grade level. To be able to handle this type of reporting problem, firms may implement data warehouses that handle what is called the "slowly changing dimension" issue.

To prevent persons who only need to query and report transaction processing system data from having any access whatsoever to transaction processing system databases and logic used to maintain those databases

The concern here is security. For example, data warehousing may be interesting to firms that want to allow report and querying only over the Internet.

Some firms implement data warehousing for all the reasons cited. Some firm implement data warehousing for only one of the reasons cited.

The need for Data Warehousing is mainly caused by the limitations of transaction processing systems. These limitations of transaction processing systems are not, however, inherent. That is, the limitations will not be in every implementation of a transaction processing system. Also, the limitations of transaction processing systems will vary in how crippling they are.

Finally, to repeat the point made initially, a firm that expects to get business intelligence, better decision making, and closeness to its clients and competitive advantage simply by plopping down a data warehouse is in for a surprise. Obtaining these next order benefits requires firms to figure out, usually by trial and error, how to change business practices to best use the data warehouse and then to change their business practices. And that can be harder than implementing a data warehouse. (www.dwinfocenter.org)

2.1.1 Advantages and disadvantages to using a Data Warehouse

There are a large number of obvious advantages involved with using a Data Warehouse. As the name suggests, a Data Warehouse is a computerized warehouse in which information is stored.

The organization that owns this information can analyze it in order to find historical patterns or connections that can allow them to make important business decisions. In this article I will go over some of the advantages and disadvantages that are connected to data warehouses.

One of the best advantages to using a data warehouse is that users will be able to access a large amount of information. This information can be used to solve a large number of problems, and it can also be used to increase the profits of a company. Not only are users able to have access to a large amount of information, but this data is also consistent. It is relevant and organized in an efficient manner. While it will assist a company in increasing its profits, the cost of computing will greatly be reduced. One powerful feature of Data Warehouses is that data from different locations can be combined in one location.

There are a number of reasons why this is important. When data is taken from multiple sources and placed in a centralized location, an organization can analyze it in a way that may allow them to come up with different solutions than they would if they looked at the data separately. Data mining is connected to data warehouses, and neural networks or computer algorithms are responsible. When data is analyzed from multiple sources, patterns and connections can be discovered which would not be found otherwise. Another advantage of data warehouses is that they can create a structure which will allow changes within the stored data to be transferred back to operational systems.

However there are a number of disadvantages that need to be mentioned as well. Before data can be stored within the warehouse, it must be cleaned, loaded, or extracted. This is a process that can take a long period of time. There may also be issues with compatibility. For example, a new transaction system may not work with systems that are already being used. Users who will be working with the data warehouse must be trained to use it. If they are not trained properly, they may choose not to work within the data warehouse. If the Data Warehouse can be accessed via the internet, this could lead to a large number of security problems.

Another problem with the data warehouse is that it is difficult to maintain. Any organization that is considering using a Data Warehouse must decide if the benefits outweigh the costs. Once you have paid for the data warehouse, you will still need to pay for the cost of maintenance over time. The costs involved with this must always be taken into consideration. When it comes to storing information, there are two techniques which are used. The first is called the dimensional technique. When the dimensional technique is used, information will be stored within the data warehouse as facts. These facts will take the form of either text or numerical information.

Data which is stored with the dimensional technique will contain information which is specific to one event. The dimensional technique is useful for workers who have a limited amount of information technology skills. It makes the data easy for them to study and understand. In addition to this, data warehouses that use the dimensional technique tend to operate quickly. The biggest problem with the dimensional technique is if the organization decides to change the way it conducts business, it will be difficult to change the data warehouse to support it. The second technique that is used storing data is called database normalisation. With this technique, the data is store in a third normal form. While adding data is easy, producing reports can be tedious Data warehouses are an important tool today, and more companies will begin using them in the future. Companies that use data warehouses need to have detailed information that is related to the transactions that take place within the group. When an organization is able to analyze detailed information that is related to its own operations, the organization can make strategic decisions.(Lester, 2003)

2.1.2 How to manage current and historical information within your Data Warehouse

In order for a company to use a Data Warehouse successfully, it must be designed so that users are able to analyze historical and current information. There are a number of things that will result from this technique, and they will have an effect on the data model design and Extracting, Transforming, and Loading (ETL) functions

As an example, a finance company may need to analyze their profits for the last three years. Looking at the data from the last three years will allow the user to view the transformations the company has gone through. Unfortunately, only having a current view will not allow the user to get the information they need.

One of the biggest challenges that data warehouse managers face today is the issue of how to manage dimensional tables over a given period of time. Not only must they be able to do this, but they must also be able to manage this information with current data. (David 1997)

2.1.3 Understanding the challenges of using Data Warehouses

While Data Warehouses can be greatly beneficial to the companies that use them, there are many challenges that a company will face in their implementation and utilization. Some experts have even said that data warehouses are one of most overrated tools in the computer industry

Many companies decide to use Data Warehouses because they simply think that it is the "next big thing," and they don't take the time to think about the requirements they will need to meet in order to use this tools.

Being able to afford a data warehouse is just one of the many requirements a company will want to look at when implementing them. Many of the companies that implement data warehouses are disappointed with its performance.

A large number of these companies fall victim to vendors that promise to help them easily implement and use the product. However, many of these vendors are more concerned with turning a profit than helping the company succeed. To face the challenge of implementing a data warehouse, it is first important for you to understand why they can be challenging. First, constructing a data warehouse is much different than constructing an OLTP system. Many companies compare data warehouses to standard OLTP systems, and this is a grave mistake. Data Warehouses are much more complex than OLTP systems. There are a large number of tools in the standard warehouse, and these tools are further broken down into many categories.

It is also important for companies to realize that data warehouses are not core business tools. What this means is that a data warehouse is much more vulnerable to the politics that may occur within a company or organization. If the data warehouse does not have the support of the employees, it will fail. Many employees have a hard time using data warehouses because of their complexity, and the companies they work for will often make the situation worse by failing to educate them. It is also challenging for companies to keep their data warehouses in tune with their production units. (David 1997)

It should be noted that Data Warehouse projects fail frequently. While they are great for the companies that properly implement them, they are disasters for companies who are not prepared. Much of the literature that is written on data warehouses is too positive toward the topic, and they don't spend enough time talking about the negatives of using a data warehouse. The biggest mistake that a company can make is not properly analyzing a data warehouse project before paying for it. By not doing detailed requirements analysis, a company sets themselves up for failure. They will waste time, and money.

2.1.4 Data Warehouse (DW) Database and Online Transaction Processing (OLTP)

The data warehouse and the OLTP data base are both relational databases. However, the objectives of both these databases are different.

The OLTP Database records transactions in real time and aims to automate clerical data entry processes of a business entity. Addition, modification and deletion of data in the OLTP database is essential and the semantics of the application used in the front end impact on the organization of the data in the database.

The data warehouse on the other hand does not cater to real time operational requirements of the enterprise. It is more a storehouse of current and historical data and may also contain data extracted from external data sources.

However, the data warehouse supports OLTP system by providing a place for the latter to offload data as it accumulates and by providing services which would otherwise degrade the performance of the database. (www.exforsys.com)

2.1.5 How to protect your Data Warehouse

While many Data Warehouses are used to access and analyze information, some companies wish to place limits on the type of information that a worker can access. While there are advantages to doing this, there are also some disadvantages as well.

Generally, an organization will not place an emphasis on security until after the Data Warehouse has been built. If most of the people who use the data warehouse will only be looking at basic reports, you will want to set up a security system that can accommodate it. While you will want to take some security measures, it is important to make sure you don't add too much. When you set up a security system for your data warehouse, there are three areas you will want to pay attention to, and these are analytical, consolidation, and standardized reporting.

The analytical aspect of Data Warehouses is the thing you will hear about the most. It is especially important during the planning stages. Despite this, over 70 percent of data warehouses are built with standardized reporting in mind. There will only be a handful of people in most companies that will know how to make plans based on the information they obtained in a analytical warehouse warehouses.

Trying to use advanced security measures in an analytical Data Warehouse is generally worthless. However, standardized reporting is a different issue. With standardized reporting, a security system is not an option. The reason for this is because it is this area of the warehouse that will have the most activity, and it is the most vulnerable to performance problems. A Data Warehouse that places an emphasis on consolidation seeks to integrate the information that it contains. Despite this, some companies may choose to merge the data into a single source. When multiple sources of information are brought together, security will become a complex issue.

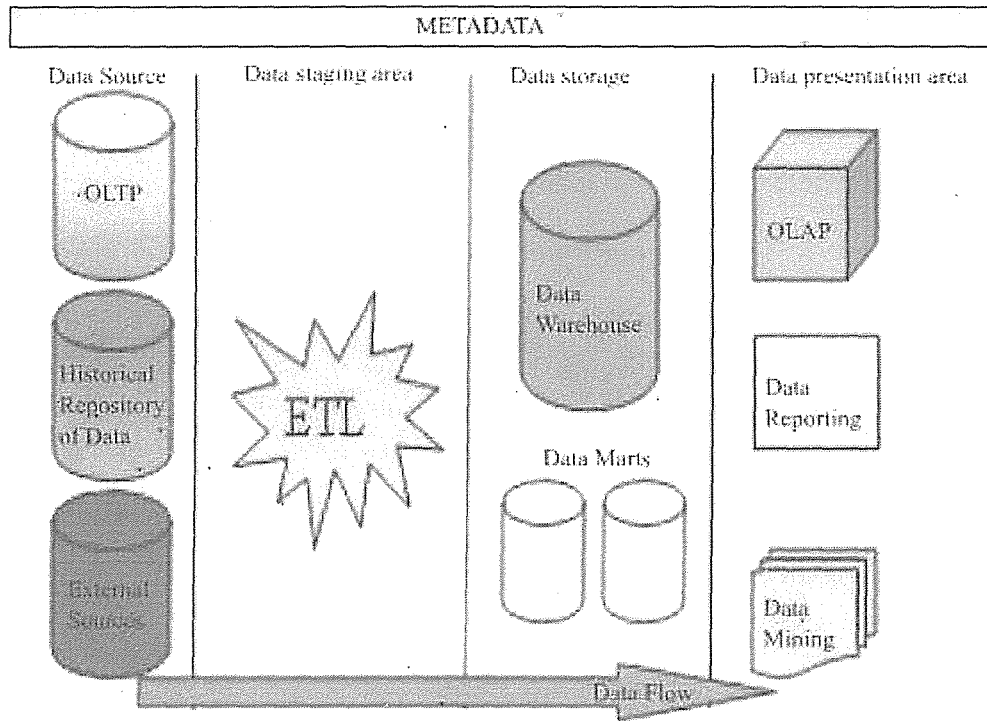
As you can imagine, the financial information within a Data Warehouse will need to have a different level of security than the information that is related to inventory. In addition to this, different departments within a company may have their own levels of security. In order for you to secure your data warehouse, it is important to make sure that each form of information has its own security system. If you are thinking of adding security to your Data Warehouse, you will need to decide where the security system should be added. The two places which are commonly chosen are the database level and application.

Placing a security system within an application is efficient because it can be connected to the data that is being processed by the application. In addition to this, the actual functions of the program can be secured as well. After the application, the next best place to add security is the data warehouse itself. When security is added to the data warehouse, all the computer programs will be secure. The one that you choose will be dependent on a number of factors. If you are using more than one program within the database, it may be best to use a database level security program. You may also want to use a database level security system if more than 100 users will be accessing the data warehouse.

Another thing that you will want to become familiar with is the security table. The security table will contain characteristics that are secured along with the identification of the user. The table will hold values which are related to the information that the user is allowed to access. (Larry, 2001)

2.1.6 Objectives of a Data warehouse and Data flow

The primary objective of Data Warehousing is to provide a consolidated, flexible meaningful data repository to the end user for reporting and analysis. All other objectives of Data warehousing are derived from this primary objective. The data flow in the warehouse also is determined by the objectives of data warehousing.



Source: [Www. exforsys.com](http://www.exforsys.com)

The data in a Data Warehouse is extracted from a variety of sources. OLTP databases, historical repositories and external data sources offload their data into the data warehouse. Achieving a constant and efficient connection to the data source is one of the objectives of data warehousing. This process is known as Data Source Interaction.([Www. exforsys.com](http://www.exforsys.com))

The data extracted from diverse sources will have to be checked for integrity and will have to be cleaned and then loaded into the warehouse for meaningful analysis. Therefore, harnessing efficient data cleaning and loading technologies (ETL—Extraction, Transformation and Loading) to the warehousing system will be another objective of the data warehouse. This

process is known as Data Transformation service or Data preparation and staging.(Www.exforsys.com)

The cleaned and stored data will have to be partitioned, summarized and stored for efficient query and analysis. Creating of subject oriented data marts, dimensional models of data and use of data mining technologies would follow, as the next objective of data warehousing. This process is called Data Storage. (Www.exforsys.com)

Finally tools necessary for query, analysis and reporting on data would have to be built into the system to the process to deliver a rich end user experience. This process is known as Data Presentation.

Users need to understand what rules applied while cleaning and transforming data before storage. This information needs to be stored separately in a relational database called Metadata.

Metadata is “data about data”. Mapping rules and the maps between the data sources and the warehouse; Translation, transformation and cleaning rules; date and time stamps, system of origin, type of filtering, matching; Pre-calculated or derived fields and rules thereof are all stored in this database. In addition the metadata database contains a description of the data in the Data Warehouse; the navigation paths and rules for browsing the data in the data warehouse; the data directory; the list of pre-designed and built in queries available to the users. (Www.exforsys.com)

2. 1. 7 Modeling Multidimensionality

According to Rudra (2002), granularity of data modeled in multidimensional data structures is an important factor for a data warehouse. Grain sizes and number of dimensions participating in the model are critical in ascertaining the quality of analytical queries that are run on such data warehouses

The most common representation of the Online Analytical Processing (OLAP) data model is a cube because this visualization is the easiest for end users to understand. Each side of the cube represents a dimension of the data being analyzed. The cells of the cube contain the facts

relating to the intersection of the dimensions. Table 1 is a simple analysis of grade distributions for courses over a number of semesters.

Table 1 Grade Distribution table

Course	Semester	A	B	C	D	F	W
MATH 0700	00F	53	69	23	14	8	8
	01S	48	58	31	12	11	4
	01U	24	10	12	4	4	4
	01F	67	72	61	10	17	1
MATH 0800	00F	30	44	28	12	5	6
	01S	48	24	13	16	7	4
	01U	16	25	27	3	8	3
	01F	48	60	38	14	6	15
ENGLISH 1100	00F	69	40	18	5	4	5
	01S	31	26	28	15	2	7
	01U	22	24	7	4	3	2
	01F	76	51	9	9	6	3

Source: Heritage International School

The logical choice for a physical database in which to store this multidimensional model would seem to be the multidimensional database (MDDB). Indeed, multidimensional database (MDDBs) produce faster query results and provide a logical grouping of data elements and their hierarchies that are easier for end users to understand. However, MDDBs suffer from several significant problems that can make them impractical for an entire data warehouse. First, cells are reserved at every intersection of the dimensions, but there may not be data to fill them, resulting in a sparse Database (Kelly 1998). For example, if we replaced the Grades dimension with a Student dimension, and stored student grades in the cells very few of the cells would actually contain data, because a single student would only be enrolled in very few of the courses offered by the school during a single semester. In reality,

the data would be summarized somewhat, before being stored in the multidimensional database (MDDDB). The disadvantage to this solution is that we have limited some of the options of rearranging the data to expand the types of analysis that can be performed.

Secondly, most multidimensional database (MDDDB's) have a limit to the number of cells and/or dimensions that can be managed (Kelly 1998). Further, the time dimension, which is always required in a data warehouse, can add orders of magnitude of complexity to the database, particularly if many different measures of periodicity are required. Multidimensional databases (MDDDBs) can be useful for storing the external views of data marts, where requirements are well-known in advance and users will not be performing frequent drill-down operations. However, the relational database structure provides the flexibility and scalability required for most data warehouses. The problem, then, is how to model the multidimensional views understood by the users in two dimensions.

2.1.8 School Information System

School information System (SIS) is a large database system which can be used for managing your school's day to day business. SIS allows users to store almost all of their school's information electronically, including information on students, employees, properties, teaching meteorites etc. Most importantly, this information can be easily shared with authorized users, records can be easily searched, and reports can be easily generated. (Ialda technology 2006-2009)

SIS is configurable and can be configured to meet most individual school's needs. It is a multi-user system and can be used by hundreds or even thousands users at same time General speaking, it is platform running on a Local Area Network (LAN). However, if the SIS server is configured to be a public server - with a static IP or domain name - it would not have a boundary limitation. Wherever you are, once you have an Internet connection and SIS client installed, you can logon to the SIS server easily just the same as if you were Sitting in the school office. However, the speed limitation is up to both your clients and server's Internet speed. SIS could make your school staff's life easier than ever. Using SIS, finding student information is just a few seconds away which might have cost hours, or even days, before. At the end of the semester, printing students' statement becomes just a few minutes' job (the speed

limitation determined by your printer), but it could be a nightmare without using SIS. If a student is absent, an email or even a short text message could be sent automatically to their parents' email address or mobile phone. (Ialda technology 2006-2009)

2.1.9 Why need SIS?

If you want to search a student's record in 2 seconds, you need SIS.

If you want to know which classroom is used by which class, you need SIS.

If you want to know who is teaching whom, you need SIS.

If you want to print students' semester statement under a minute, you need SIS.

If you want to print new semester information letter under a minute, you need SIS.

If you want to view a class timetable in less than 5 seconds, you need SIS.

If you want to know where your properties are and how much they're worth, you need SIS.

If you want to automatically send emails or short text messages to students' parents if something happens, you need SIS.

And many more ... (Ialda technology 2006-2009)

2.2.0 More information on school Information system

The parents of students are very busy now days, this school management system helps the parents monitor their children from anywhere. They can check their children's academic performance from a remote location, plus the teacher's can post specific notes about the student to their parents. Using this school Information system the users can see the attendance in various formats, day-wise, monthly attendance, session or total attendance, depending on the

permissions the user has. Marks or report card can be viewed or added, report cards generated, merit list, highest marks etc.

2.1.1 Data Warehousing Integrates School Information Systems.

Using a Data Warehousing system to act as the backbone for connecting a school's disparate computer systems may help the school run better, a presenter said at the session, Data Warehousing: Making a Difference in Improving Schools.

Bill Flaherty, director of technology services at Hanover County, Va., Public Schools, said data warehousing can improve the way student data is used in schools.

Flaherty said data warehousing is simply a database structure that has access to the varying school systems, like enrollment, library, grades and food services.

(<http://findarticles.com/p/articles>)

2.1.2 Data Models for Data Warehouse

Data models for Data warehouses, also called Online Analytical Processing (OLAP) systems, are considerably simpler, because the intent is for end users to interact with the models to produce their own reports. Creating these simplified models requires careful analysis of data sources in order to discover data that will be most useful for analysis and then organizing that data so that it is easy to access and understand (Lester, 2003).

The purpose of having a good design is to have a product at the end that will be of benefit to the organization. The best judge whether or not a data warehouse is 'good' is the end user. (Lester, 2003).

Therefore, the best way to evaluate the quality of the Data Warehouse is through customer satisfaction surveys. To develop these surveys, we need to understand the factors that measure the quality of Data Warehouse and the components of the warehouse that impact those factors. (Lester, 2003).

The foundations of Data Warehouse Quality Project (DWQ), a 3 – year ESPRIT research project (Vassilious, 1997), set out to develop a quality that is specifically mapped to the Data Warehouse.

Figure below shows the five quality factors that we used as the framework for DWQ, and the components of the Data Warehouse, In addition a set of baseline performance requirements, and metrics to measure the properties of the Data Warehouse components, must be developed and then formulated. Test to relate the baseline goals to the metrics.

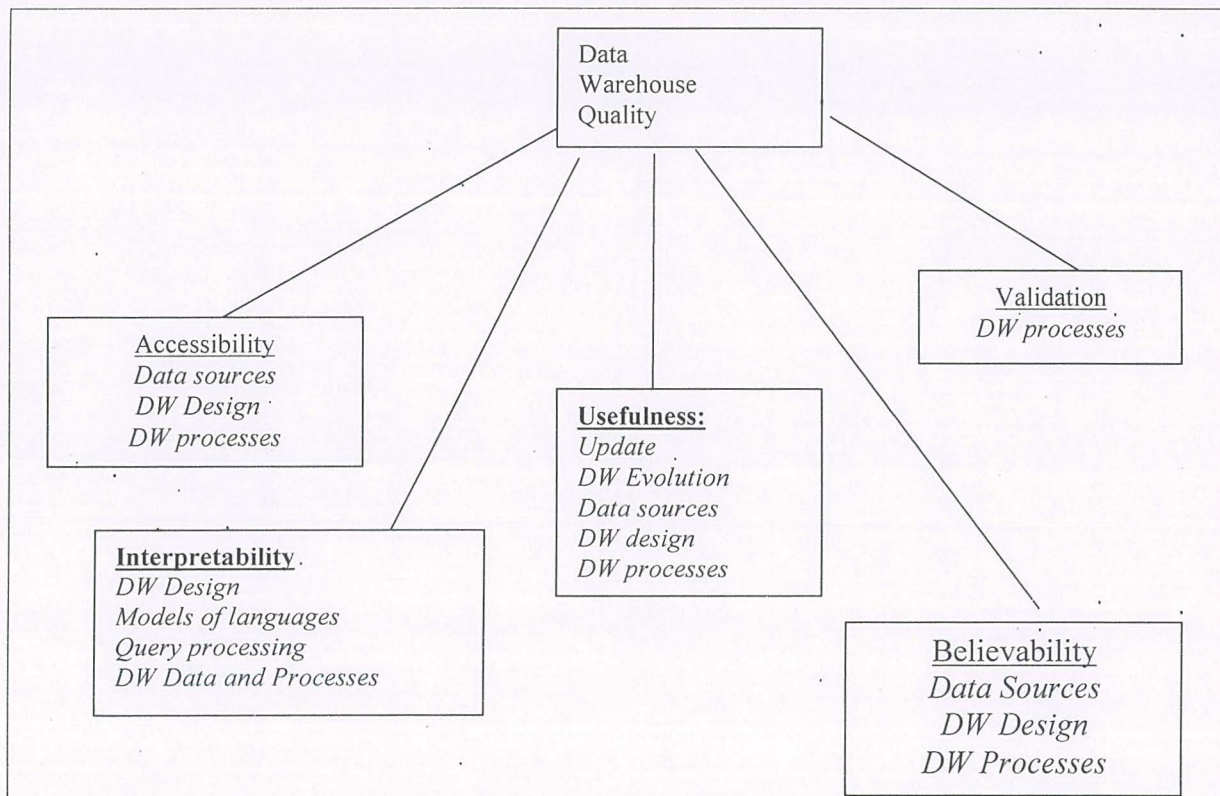


Figure 2- The DWQ quality framework (Vassilliou 1997)

2.1.4 The Methods for designing Data Warehouse Architecture

There is disagreement in the data warehouse community about the best method to use for designing the warehouse architecture. The approaches recommended generally fall into one of two categories. The first method involves meticulous analysis of the source data and then constructing models for a comprehensive data warehouse. This method will be referred to as the bottom up method. The second approach, which we will call the top-down. Method looks at potential data marts that will be needed for the overall Data Warehouse and then designs data models to support the data marts. The data warehouse is considered complete when all of the data marts have been constructed.(Lester, 2003) The details of each method will be outlined in the sections that follow.

2.2.5 The Bottom-Up Approach

The bottom-up method described in (IBM ,1998) begins with a clear understanding of the conceptual view of the business as a whole and then selects the parts of that view that will be useful for business analysis in the data warehouse. Table below outlines the six steps of the bottom-up method.

Steps:

- Create ER diagram of data source(s)
- Remove associative entities and subtypes
- Merge entities in M:N relationships
- Identify parts of the model to be used in DW
- Identify the measures and dimensions
- Create facts

The first step is to create the ER diagram of the source data, if one does not already exist. Then, all associative entities and subtype entities are removed from the diagram. Next, the

model is simplified by rolling up entities at the ends of many-to-many relationships into single entities. As entities are eliminated, any attributes associated with them that would be useful for analysis should be retained.

Once the model has been simplified, the designer works with business executives and analysts to determine what parts of the model will be valuable in the Data Warehouse, as well as identifying information that is not currently included in the model. At this point, the dimensional model can be constructed. This is done by identifying the measures that will make up the fact table and the dimensions that are desired and adding a time dimension. The next step is to create the facts. The designer can determine the desired level of detail, or granularity of data, from the requirements. One of the factors to consider in this step is the additivity of the measures. Additivity is the ability of a measure to be summarized. Some measures, such as percentages, are non-additive, as they cannot be summarized across the time dimensions to produce meaningful information. In general, all measures should be fully additive. If they are not, they may need to be broken down further. After all of the facts have been created, they are merged into a single fact table and linked to the dimension tables, completing the star schema.

2.2.6 The Top-Down Approach

The top-down design method begins by looking at the external user views of the organization by studying query behavior. The potential data marts are then identified by creating materialized views that can satisfy the queries. A materialized view is defined as a preprocessed query that is stored for faster retrieval. A carefully designed materialized view can serve many different queries. There is currently much research in the academic community involving identifying and constructing materialized views, particularly views that are self maintaining. To build a data mart, materialized views are analyzed to discover the common facts and dimensions being used in the queries (Kimball, 1999). The next step is to construct a matrix of the potential data marts (fact tables) and the dimensions they would have in common (Moriarty, 1995).

At this point, it may be necessary to have user groups meet to resolve any conflicts in business rules or definitions that might appear in the dimension tables. For example, the Admissions and Records department will view data primarily by academic calendars, while the Business Office will view their data on a fiscal year basis. In such situations, a compromise must be reached,

whether it is a common calendar, or an agreement that both calendars will be represented in the dimension table. Once all conflicts have been resolved, the star schemas are designed using the matrix. Finally, the source data are analyzed to locate the information that will be needed to populate the data marts. (Lester, 2003)

CHAPTER THREE

RESEARCH METHODOLOGY

This section describes the tools and methods used in carrying out this research.

3.1.0 Research Instruments

The research employed a number of instruments; questionnaires, observation, and interviews

Questionnaire was used because it is easy to administer and time saving. The response given is also easy to analyze.

Observation was also employed especially in the current Data Warehouse for Heritage International School. This instrument was important because it helped confirm the answers in the questionnaire whether they are true or false.

Lastly, the interview instrument was also employed. It involved face-to-face interaction with the staffs and students of Heritage International School.

3.1.1 Developing a Survey

In order to develop a baseline set of end user expectations, a survey was developed to determine the features or characteristics of a data warehouse that would be required for the intended users to consider the data warehouse a successful project. Using the Data Warehouse Quality (DWQ) quality framework. Vassiliou (1997), as a guide, first, the quality factors had to be defined, then metrics determined for those qualities, and, finally, questions developed to be asked of the end users in order to establish a range of acceptable values. The descriptive survey was applied. It included both primary and secondary data where primary data was gathered through interviews and observation, while secondary data was gathered through documentary study.

3.1.2 Study area

The study focused on developing an Information System for Heritage International School. Heritage International School is located about 2.5km from Kampala city centre and 1km off Ggaba road along Kiwafu road. The School has a population of about 200 students and 40 staffs.

3.1.3 Procedure of the study

Permission was obtained from both the school chief administrator of Heritage international school and systems information administrator, Heritage International School to assess and analyze the information system at the school. The study applied different basic and appropriate methods in the study; this including among others, observation, direct study of the database, interview all the concerned stake holders at the school.

3.1.4 Data collection methods

The data was collected by using questionnaires, interviews, observation, and school documents.

In developing this system, the usual system, development life cycle was observed. These Include System Analysis, System Design, Coding, Testing, and Implementation.

CHAPTER FOUR

SYSTEM ANALYSIS AND DESIGN

4.1.0 Description of the Current System

The current School Information System allows users to store almost all of their school's information electronically, including information on students, employees, properties, teaching meteorites etc. It also

enables the students, teachers, parents and the management stay in touch with each other at all times. However the system does not enable validation of data entry to enable consistency and accuracy, and does not produce the required reports,. It also does not meet test score and assessment information sometimes, and it doesn't enable one to store everything from activities, assessments, and to other mandated data.

4.1.2 Strengths of the current system

Following the analysis of the current system, the following were the strengths of the current system;

The system in existence has got the advantage of having an intact record system available in the hard disc.

The current system is cheap to expand incase need for more data capture.

The current system has got its data normalized. This is instrumental in having an unduplicated database for the school.

The information system setup is not so much complicated to require skilled experts to run the system.

There is also no much resource was needed to setup the whole system. For example computers, network cables, switches, furniture, and medium skilled personals to run the system.

4.1.3 Weakness of the current system

An investigation was carried out to establish how the existing School Information System functions and what its problems were. The current School Information System enables the students, teachers, parents and the management be in touch with each other at all times.

Howbeit, according to the school chief administrator, the system does not enable validation of data entry to enable consistency and accuracy.

It also does not meet test score and assessment information sometimes.

The current system doesn't enable one to store everything from activities, assessments, enrolment, and to other mandated data.

The system lacked periodic maintenance and backup of the data. This could to the overall loss of the schools' data incase of a major catastrophe like fire.

The study also discovered that the security of data was data was too poor. For instance, the server room has got a weak door that could render easy access of trespassers into the server room. The server password is also not kept secret by the systems administrator. The server password is known to other top school managements.

4.1.4 Possible solutions to the problems of the current system

- There is need for automated computer based information system to enable easy access, retrieval, and validation of data.
- A proper Data Warehouse should be designed and developed to enable easy access on the past and present information of the School.
- There is great need for a well trained Database Administrator to handle the School Database.

4.1.5 The Proposed System

The new system is supposed to cater for the school's data security both physically and logically.

The new system needs to have a wide data capture capacity to ensure that all the schools data is captured in the information system.

The system should have a periodic maintenance time table in place. This should be timely followed to ensure the smooth running of the system.

4.1.6 Requirements of the proposed system

- The system after completion is expected to Process and provide Heritage International School with information about the students in the school and outside the school hence providing the school administrators with meaning report in decision making.
- To provide Student Information System access to the teachers, the Principal, the Parent, and the students (Data entry, update, and deletion)
- To document the Student Information System within the standards and procedures manuals.
- To provide training to all the users responsible for the Student Information System and the Data Warehouse.
- Keep tract of the students, staffs, assets, and departments..
- Keep track of all the staffs/teachers in Heritage International School.
- Improve employee service and satisfaction level through faster response times.
- Provide central Information which can be used to attend to a problem requests, provide detailed information that can be used for decision making quickly and accurately.
- Produce timely professional looking and summarized reports that can be used for decision making.
- Provide a good interface not only for expert user but also for the novice user who knows little about the retrieval mechanism.

4.1.7 System Design

Based on the user requirements and the detailed analysis of a new system, the new system was designed. This is the phase of system designing. It is a most crucial phase in the development of a system. Normally, the design proceeds in two stages:

- **Conceptual design**

Conceptual Database design involves modeling the collected information at a high-level of abstraction without using a particular data model or DBMS (Database Management System).

Below is a Data Flow Diagram showing how data flows through Heritage School Information System. It shows how data enters and leaves the system.

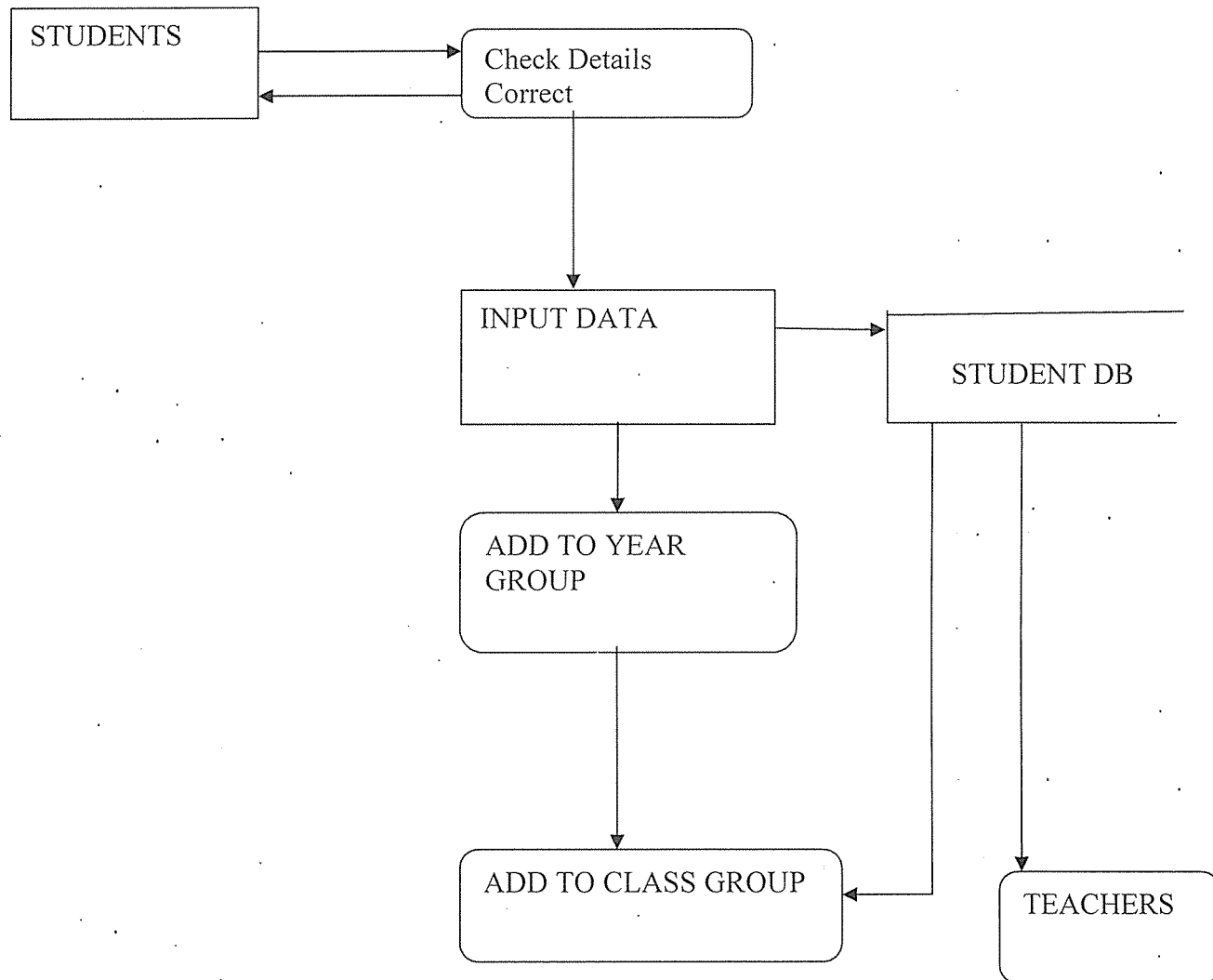


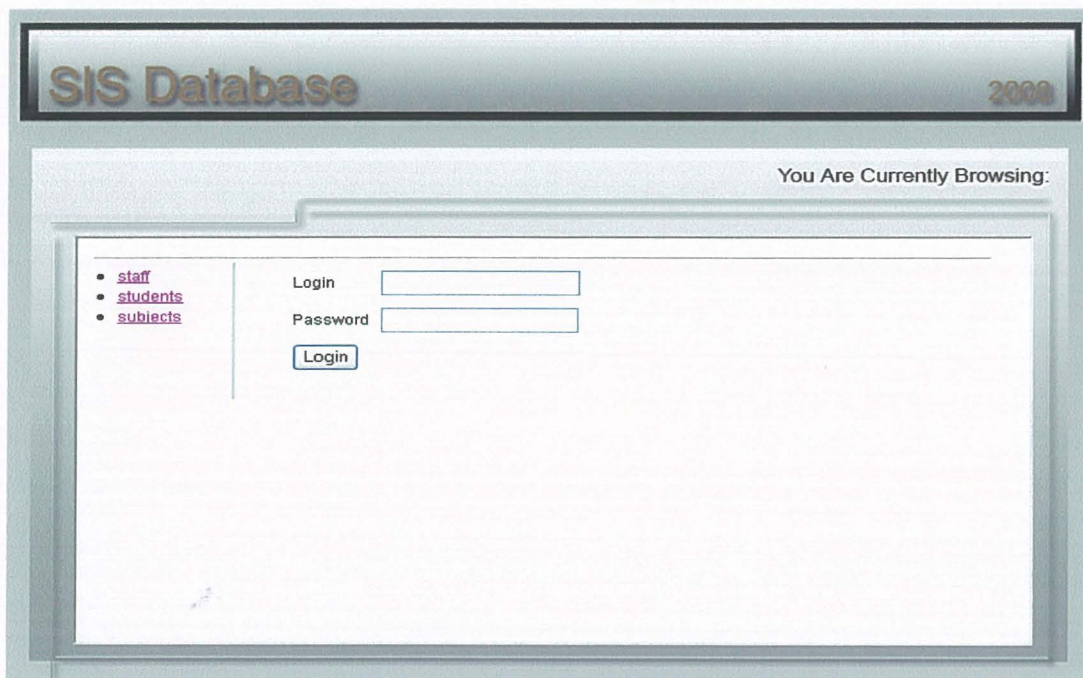
Figure 3: Data Flow Diagram

Once checked the data can now be entered in to the students database the students must also be entered to a year group as a part of this last process the students will get to choose which subject they should to do.

4.1.8 Physical design

Interface Design describes the input and output interfaces of the system

User Login Interface



The image shows a web-based user login interface for a system titled "SIS Database". The interface has a header bar with the title "SIS Database" on the left and the year "2009" on the right. Below the header, there is a section titled "You Are Currently Browsing:". On the left side of this section, there is a list of links: "staff", "students", and "subjects", each preceded by a bullet point. To the right of this list, there is a login form with two input fields: "Login" and "Password". Below these fields is a "Login" button.

SIS Database 2009

You Are Currently Browsing:

- [staff](#)
- [students](#)
- [subjects](#)

Login

Password

Login

The user will login by typing in his/user name.

Output Interface: Students

SIS Database

2009

You Are Currently Browsing:

- [assets](#)
- [students](#)
- [subjects](#)

[\[Logout\]](#)

Staff

Records shown 1 - 1 of 1

[Add Record](#)

			staff_id	f_name	l_name	department	course
View	Edit	Delete	BM001	Muhamed	Ssendagire	Business and Management	Economics

Copyright 2008

Output Interface: Assets

SIS Database

2009

You Are Currently Browsing:

- [assets](#)
- [staff](#)
- [subjects](#)

[\[Logout\]](#)

Students at School

Records shown 1 - 2 of 2

[Add Record](#)

			student_id	residence	department	course
View	Edit	Delete	BBA001	Kansanga	Business and Management	BBA
View	Edit	Delete	BCS21111	Kansanga	Computer Science	BCS

Copyright 2008

Output interface: Staff

The screenshot displays a web application titled "SIS Database" with a copyright notice of 2008. The main heading is "You Are Currently Browsing:". On the left, a sidebar lists navigation options: [staff](#), [students](#), and [subjects](#). The main content area is titled "Index Page" and features a form for adding or updating a record. The form includes four input fields: "asset_id" (containing "T001"), "asset_class" (containing "Furniture"), "department" (containing "Procurement"), and "asset_description" (containing "Small Table, Light Brow"). A "Post" button is located below the form. A "[Logout]" link is positioned in the top right corner of the main content area. The footer of the page reads "Copyright 2008".

asset_id	T001
asset_class	Furniture
department	Procurement
asset_description	Small Table, Light Brow

4.1.9 Database

Database is a structured collection of records or data that is stored in a computer system.

Below are the sample Database tables containing Student, Enrollment, courses, and the teachers

for Heritage International School.

Students : Table		
	StudentID	StudentName
▶	+ S1	George Smith
	+ S2	Alex Bamford
	+ S3	Jessica Peters
	+ S4	Mary Lai
	+ S5	Rashid Ahmed
	+ S6	Daniel Arnold
	+ S7	Holly Bowers
	+ S8	Sarah Turner
*		
Record:	1	

Enrolments2 : Table		
	StudentID	CourseID
▶	S1	BS
	S1	MM
	S2	BS
	S3	EN
	S3	HI
	S3	MA
	S4	EN
	S4	HD
	S5	EN
	S6	EN
	S6	MM
	S7	BS
	S7	EN
	S7	HI
	S8	MA
*		
Record:	1	

Courses1 : Table			
	CourseID	Coursename	TeacherID
	BS	Business Studies	T5
	EN	English	T1
	HD	Hairdressing	T4
	HI	History	T3
	MA	Mathematics	T5
	MM	Motor Mechanics	T2
▶			
Record:	7		of 7

Teachers : Table		
	TeacherID	TeacherName
	T1	Eliza Newton
	T2	Michael Ford
	T3	Norman Castle
	T4	Cynthia Lau
	T5	Michelle Davies
▶		
Record:	6	

Student table

Fname	Lname	Regno	ID	DOB	ADDRESS	Parents Name	Width
X(10)	X(10)	A(5)	N(3)	Dd/mm/yy	X(20)	X(20)	16

Staff Table

Fname	Lname	ID	Title	Department	Total
X(10)	X(10)	N(3)	X(3)	X(5)	N(10)

Department

DeptName	Dept ID	DeptNum	Location	Type
X(6)	N(3)	N(3)	X(6)	X(12)

Subject/Course Table

CourseName	CourseID	Type	Scores
X(20)	N(3)	X(6)	N(3)

4.2.0 An entity-relationship (ER) diagram

An Entity-Relationship (ER) diagram is a specialized graphic that illustrates the interrelationships between entities in a database. ER diagrams often use symbols to represent three different types of information. Boxes are commonly used to represent entities. Diamonds are normally used to represent relationships and ovals are used to represent attributes. Below is an illustration showing relationships in a school information system for Heritage International School.

An E- R Diagram for teaching staffs (Professors) at Heritage International School

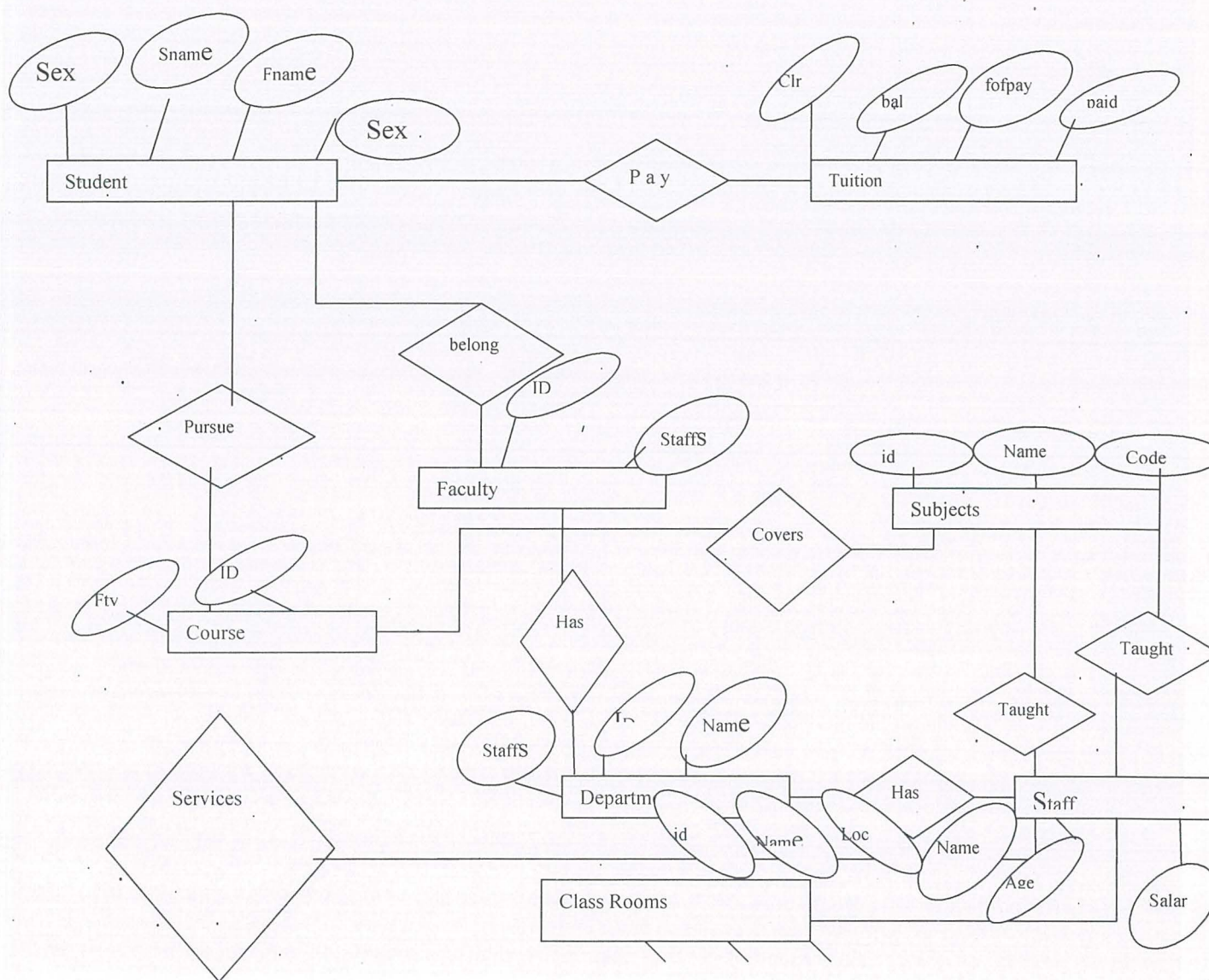
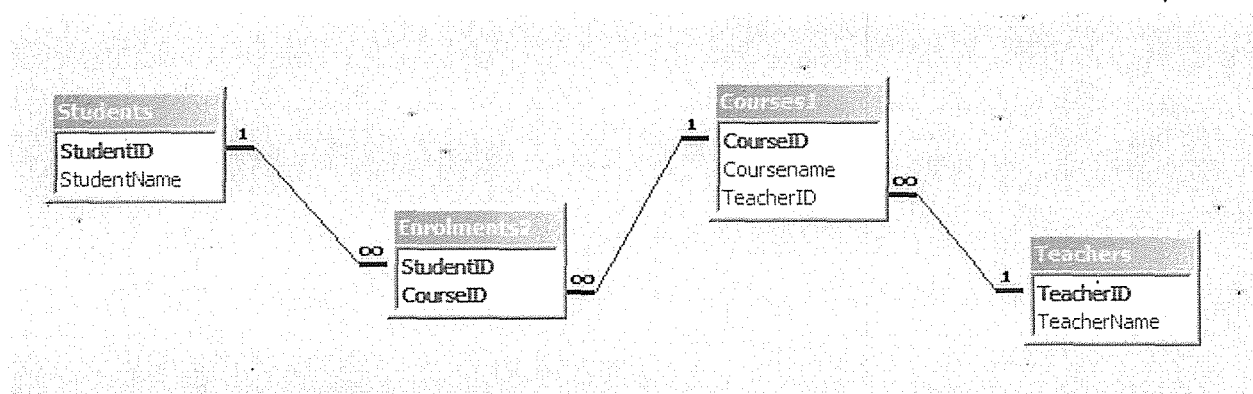


Figure 4: E-R Diagram for a School Database

The ER diagram tells us which tables should be created. It also tells us about the relationships between the tables. An ER diagram will make database (SIS) design a lot easier

Remember that a many-many relationship between two tables or entities must be resolved into two one-many relationships - this is done by taking the primary key of each table or entity and combining them into a composite primary key for a new table.

The diagram tells us which tables we should create: Students, Teachers, Courses, and Enrolments. The relationship diagram for the tables looks like this:



Note that the Enrolments table (which was created to resolve a many-many relationship) has a composite primary key, consisting of the primary key from the Students table and the primary key from the Courses table.

4.2.1 Designing the Star Schema

The star schema (sometimes referenced as star join schema) is the simplest style of data warehouse schema, consisting of a few "fact tables" (possibly only one, justifying the name) referencing any number of "dimension tables". The "facts" that the data warehouse helps analyze are classified along different "dimensions": the fact tables hold the main data, while the usually smaller dimension tables describe each value of a dimension and can be joined to fact tables as needed. (Lester, 2003)

Dimension tables have a simple primary key, while fact tables have a compound primary key consisting of the aggregate of relevant dimension keys.

It is common for dimension tables to consolidate redundant data and be in second normal form, while fact tables are usually in third normal form because all data depend on either one dimension or all of them, not on combinations of a few dimensions.

The star schema is a way to implement multi-dimensional database (MDDB) functionality using a mainstream relational database: given the typical commitment to relational databases of most organizations, a specialized multidimensional DBMS is likely to be both expensive and inconvenient. Another reason for using a star schema is its simplicity from the users' point of view: queries are never complex because the only joins and conditions involve a fact table and a single level of dimension tables, without the indirect dependencies to other tables that are possible in a better normalized snowflake schema. (en.wikipedia.org/wiki/Star_schema) The first step in the process of designing the star schema for the data warehouse is to identify the measures and dimensions. For the measures, we need to determine the level of detail, or granularity, of data that we need that will provide the most flexibility. The reports that we will be producing actually use a variety of measures, so the objective is to define those measures and determine how they can fit into one or more fact tables. (Lester, 2003)

The figure shows Heritage International School database schema

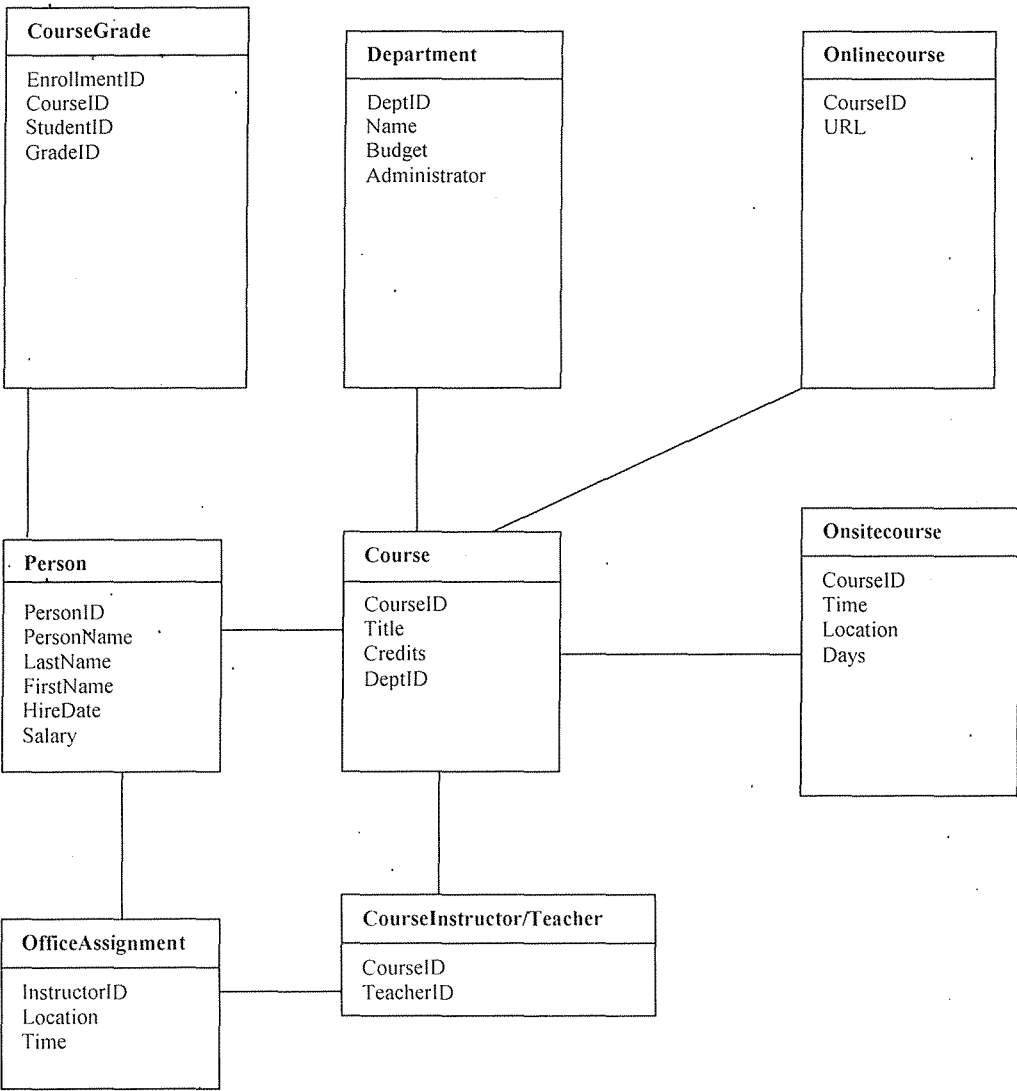


Figure 5: Database Schema

CHAPTER FIVE

SYSTEM IMPLEMENTATION AND EVALUATION

5.1.0 Program Implementation

Having carried out the analysis of the current system and designed the proposed new system, this section is focused on the implementation of the new system. The new system will be implemented using PHP, Microsoft Access, and MySQL.

Different graphical user interface forms, queries and reports were designed. The choice of a graphical user interface was taken up because its user friendly, it does not require a lot of technical training for a user of the system.

5.1.2 Features of the new system

Form Page:

The form page enables the user to easily enter and retrieve data.

Report Page

A report is an effective way to present your data in a printed format. Because you have control over the size and appearance of everything on a report, you can display the information the way you want to see it.

Query Page

Enables the user to manipulate the Database

Tables

Stores all information.

CHAPTER FIVE

SYSTEM IMPLEMENTATION AND EVALUATION

5.1.0 Program Implementation

Having carried out the analysis of the current system and designed the proposed new system, this section is focused on the implementation of the new system. The new system will be implemented using PHP, Microsoft Access, and MySQL.

Different graphical user interface forms, queries and reports were designed . The choice of a graphical user interface was taken up because its user friendly, it does not require a lot of technical training for a user of the system.

5.1.2 Features of the new system

Form Page:

The form page enables the user to easily enter and retrieve data.

Report Page

A report is an effective way to present your data in a printed format. Because you have control over the size and appearance of everything on a report, you can display the information the way you want to see it.

Query Page

Enables the user to manipulate the Database

Tables

Stores all information.

5.1.3 Program Testing

Screen shot for Staffs'

ID	Fname	Lname	Age	Nationality	Stat	Address	Parents	Phone	St	Sex	Grade
0001			0					0			0

Screen shot for staffs

ID	StaffID	Name	Little
AutoNumber	0		
Qualification	Salary	Department	
	\$0.00		
Subject			

5.1.4 Sample Testing for SIS 1

ID	Fname	Lname	Age	Nationality	Stat	Address	Parents	Phone	St Sex	Grade
1	Char	Moses	10	Ug	EA	BOX 20 Kampala	Moses	+06	Male	8p
*	number		0					0		0

5.1.5 Program Flow Chart

The program flow Chart illustrates the information flow from one point to another as indicated below

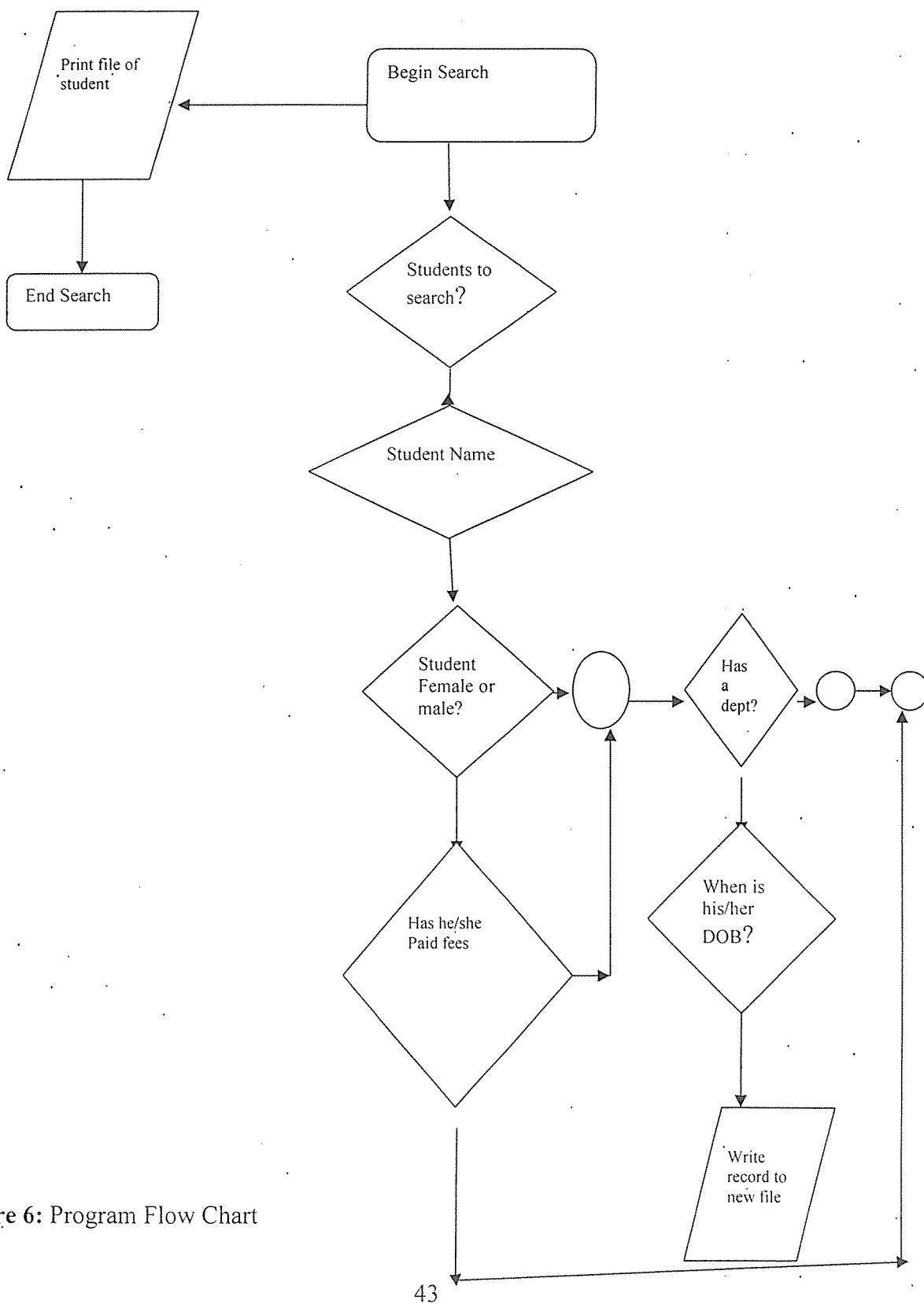


Figure 6: Program Flow Chart

5.1.6 Implementation

Installation requirements;

For optimum use of the application, the systems administrator will install the following hardware specification. The application computer needs to meet the following standards;

- At least 60 GB of hard disc space
- Minimum of 1024 MB of RAM
- 2.4 GHz of processor speed
- A gigabit Ethernet adapter
- Any printer for report printing
- A good resolution color monitor
- Local Area Network

The operation systems required during installation are the following

- Windows 98
- Windows XP
- Windows Vista
- Any Microsoft operating system

5.1.7 Change over methods

The system was implemented successfully. The system will be updated regularly and it will be performed in parallel format. It will be allowed to operate alongside the current system until the new system stabilizes and the old system may be discarded.

5.1.8 User Manual

The user of the program should follow the following steps:

- Go to Start
- All Programs
- On the program menu
- Click SIS

5.1.9 System Evaluation

Some of the issues encountered in designing and constructing the Database were specific to the structure of the source data. The issue of duplicate data elements, and the necessity of determining a copy of record. Several data elements are duplicated in different files, with no guarantee that all copies are kept in synch. In order to ensure that the data warehouse contains the most reliable data, the developer engaged in analysis to determine which copy of the data end users consider to be most correct. In proper, normalized, relational Databases, this is not an issue

The interactive part of the system with the users has been considered, this is by going through the implementation plan of the system. Key ethical issues of the use of the system have been mentioned. Therefore, it is very important for Heritage International School management to be aware of them.

5.2.0 Security requirements

All the users have to first login with a correct username and password to gain access to the system.

The secure password is issued by the systems administrator to any user who wants to gain access into the system. Without a secure password, no user will gain an authorized login into the system.

Below is the control measures that will be undertaken by a user to login into the system.

It is password protected to prevent unauthorized users. The user must have a correct user name and password to login into the system.

The server computer will have a strong anti virus installed to prevent malicious attacks and virus programs from interfering with the system processes.

Back up and periodic maintenance is recommended for all systems.

CHAPTER SIX

DISCUSSION, CONCLUSIONS, AND RECOMMENDATION

The purpose of the research described in this paper was to analyze the process of designing and constructing a Database for Heritage International School and to document the special characteristics of School Information system that affect the data warehousing processes. Some of these characteristics are specific to data sources organized as complex flat file structures. This chapter discusses the conclusions reached as a result of this research.

6.1.0 Top-Down vs. Bottom-Up Method

The analysis of current reporting behavior showed that the top-down method could not meet the requirements for the Database. The top-down approach makes the assumption that reports currently being produced from the source systems are correct, and the research done in this project has shown that this is not the case. Current reports frequently make use of hand-keyed data elements that represent other original elements, thereby introducing an increased possibility for error. In addition, the fact that the source systems are not currently integrated at all means that unreliable data elements must sometimes be used for reporting, even if a more reliable data element exists in another system. Once the unsuitability of the top-down method had been determined, the focus of the research shifted from comparing the two design methods to determining if the reporting requirements could be met using the bottom-up method and documenting problems encountered during the process. In the evaluation of the Database that the bottom-up method was indeed successful in meeting the requirements for the Database analysis, as well as end user performance requirements. Also, several of the end users expressed an interest in having other types of reports available from the Database, in addition to those defined for the DSP analysis. Because the bottom-up method was used, much of the data needed for those reports have already been extracted and incorporated into the existing Database.

6.1. 1 School Information System-specific Issues

One of the issues encountered specific to School Information System (SIS) that affects the Database process concerns the time characteristics of the data. In most business-oriented database applications, sales and production data are entered into the system on a continual basis. Time constraints on the data are somewhat artificial, meaning that the information is grouped into time categories for accounting convenience. By contrast, academic data are stored in snapshots, with very definite beginning and ending dates. Primarily, institutions tend to be most interested in the data as they exist at the official census date and at the end of the semester. This means that frequent refreshes of the data are not necessarily important, but the refreshes must be done promptly and completely on those critical dates.

A second characteristic of Database is that they tend to be more complex and multidimensional than business-oriented data for analytical applications. Research of star schema development for business applications implies that one well-designed star schema can meet most of the reporting needs of a particular area of the organization. In these applications, however, the Measures are fairly straightforward, and just a few dimensions exist. In some of the reports, the measure was actually the count of joins of the dimensions, while other measures were evident in the fact table. A single star schema, with all of the dimensions necessary for the reports, would have been too complicated for end users to work with, so the solution was to use more, simpler star Schemas to support each different aspect of the analysis

6.1. 2 Lessons Learned

One of the phases of the research project that proved to be the most cumbersome was the development of extraction routines. Most of these routines are programs run on the source systems, and the logic for cleansing the data is in those programs. That means that any time the cleansing rules change, the programs have to be modified and recompiled. The preferred method for extracting and cleansing data is to extract the data as is from the source systems into a staging area that has an identical structure to the relational database that will contain the clean data. All of the cleansing routines are incorporated into the transformation processes. The benefits of this method are two-fold: (1) the effect of the extraction routines on source system performance is minimized, since the overhead of the cleansing process is removed and, (2) modifications that need to be made to the cleansing or loading routines can be made on the fly.

A second item that proved to be a problem during data loads and refreshes was the definition of foreign keys in the physical database. The foreign keys in the data warehouse that were represented in the logical data models were also implemented in the physical Database. This resulted in several incomplete table loads, as records with old values for certain fields that are not longer valid (such as major codes or academic program codes) were encountered. Implementing the foreign key constraints does not really make sense for a data warehouse, since data entry control is not an issue. Most of the constraints have subsequently been removed.

Finally, as reports were developed for the end users, the importance of identifying the logical hierarchies that exist in the data became evident. Some of the star schema designs could have been done much differently to take advantage of those hierarchies and provide more flexible and informative reports.

The users were asked the minimum acceptable hours of availability each day, if they would prefer access during normal working hours over weekend and after-hours access, and how often they might expect to use the system. They were also asked how important the availability of the system is to them compared to other criteria. Questions regarding the access to data asked about current levels of access to the source system, the importance of viewing detail data, and the necessity of the ability to manipulate data themselves.

6.1.3 Future Recommendation

The results of the evaluation indicate that the users who have had the opportunity to use the Data Warehouse (DW) and School Information System (SIS) view it as a successful project. Most all of them accepted that it provided them with information they had no to access previously, in an easy to use and easy to understand format. The key to ensuring that it remains a valuable information resource to the college will be continual modification, expansion, and tuning. As mention in the previous section two the extraction routines need to be rewritten to streamline the ETL processes. In addition, more data sources should be incorporated to expand the database to meet users. Increasing need for information

BIBLIOGRAPHY

Ballou, (1999). Enhancing Data Quality in Data Warehouse Environments. Communications of the ACM. (January): 73-78.

David (1997), From a Relational to a Multi-dimensional Data Base
Internet. Available from
<http://www.essentialstrategies.com/publications/datawarehouse/relmult.htm>;
accessed 18th June 2008.

Dayal (1997), An Overview of Data Warehousing and OLAP Technology. ACM Sigmod Record. (March).

Gardner, (1998). Building the Data Warehouse. Communications of the ACM. (September): 52-60.

<http://www.exforsys.com>. Accessed on 29 Sept 2008

Inmon (1996), Building the Data Warehouse. New York. John Wiley & Sons, Inc.

Inmon (2001), Little White Lies. Internet. Available from
<http://www.billinmon.com/library/articles/artlies.asp>; accessed 5th July 2008.

IBM. 1998. Data Modeling Techniques for Data Warehousing. IBM Redbook SG24-2238-00. Armonk. IBM Corp.

[HTTP://www.iada technology](http://www.iada technology) (2006-2009). Accessed on 16 Aug 2008

Kimball. A Practical Method for Planning a Data Warehouse. Data

Web house. Internet. Available from
http://www.intelligententerprise.com/db_area/archives/1999/990712/webhouse.shtml; accessed 17 July 2008.

Lauren. (1997). Anatomy of a Failure. CIO Enterprise Magazine.
(November). Internet. Available from
http://www.cio.com/archive/enterprise/111597_data_content.html;
Accessed 30 June 2008.

<http://www.westminister.org.uk>
Accessed on 5/09/20

Larry. (2001). The Case Against Data Warehousing. (June). Internet. Available from
<http://www.dwinfocenter.org>; accessed 4th July 2008.

Laura. (2000). Developing a Data Warehouse Architecture. Internet.
Available from <http://www.users.qwest.net/~lauramh/resume/thorn.htm>;
Accessed 24 June 2008

Lester (2003), design of data warehouses.

Terry (1995), Modeling Data Warehouses. Database Programming & Design. Internet.
Available from <http://www.inastrol.com/Articles/9508.htm>; accessed 22nd July 2008.

Web house. Internet. Available from

http://www.intelligententerprise.com/db_area/archives/1999/990712/webhouse.shtml; accessed 17 July 2008.

Lauren. (1997). Anatomy of a Failure. CIO Enterprise Magazine.

(November). Internet. Available from

http://www.cio.com/archive/enterprise/111597_data_content.html;

Accessed 30 June 2008.

<http://www.westminister.org.uk>

Accessed on 5/09/20

Larry. (2001). The Case Against Data Warehousing. (June). Internet. Available from

<http://www.dwinfocenter.org>; accessed 4th July 2008.

Laura. (2000). Developing a Data Warehouse Architecture. Internet.

Available from <http://www.users.qwest.net/~lauramh/resume/thorn.htm>;

Accessed 24 June 2008

Lester (2003), design of data warehouses.

Terry (1995), Modeling Data Warehouses. Database Programming & Design. Internet.

Available from <http://www.inastrol.com/Articles/9508.htm>; accessed 22nd July 2008.

APPENDIX A

Data Warehouse Questionnaire/Survey questions

1. Rate the following items in order of importance to you.

- ☐ Number of hours the system is available.
- ☐ Days and times the system is available.
- ☐ Ability to view detail data when needed.
- ☐ Ability to download and manipulate desired data.
- ☐ Same or better access to data currently accessed from SIS.
- ☐ Other (please explain) _____

2. How would you rate the importance to you of the ability to extract raw data (i.e. data in text format) from the data warehouse to manipulate at the desktop level, for example, in Excel or Access?

- ☐ Very important
- ☐ Somewhat important
- ☐ Not important

3. For standard reports that you will be processing frequently, what is the maximum acceptable response time?

4. Indicate your current level of use of the SIS system.

- ☐ I use the system on a daily basis
- ☐ I use the system several times a week
- ☐ I occasionally use the system during the semester
- ☐ I rarely or never use the SIS system

5. In situations where data from the data warehouse conflicts with SIS reports, what would be the most compelling evidence to you that the Data Warehouse data is correct?

- ☐ The detail data that was used to derive the summary

_____ Documentation that shows the source(s) of the data

_____ Other (please explain) _____

6. Select the option that best describes your preference for producing reports.

_____ I would like an icon on my desktop that will automatically display my reports

_____ I would like to be able to select from a list of pre-defined reports

_____ I want to be able to produce my own reports using a graphical interface

_____ I want to be able to produce my own reports from a command line interface

_____ Other (please explain) _____

7. What is the minimum number of hours of up time per day that you consider acceptable?

8. Describe your current level of *viewing* access to data in SIS.

_____ All or partial access to student data

_____ All or partial access to instructor data

_____ All or partial access to course data

_____ Unlimited access

_____ Other (please explain) _____

9. Rank the following in order of your preference regarding documentation for a software application.

_____ Hardcopy user's manual

_____ Online documentation

_____ Context-sensitive help

_____ Diagrams and charts

10. If you are using a graphical interface to create your own reports, what is the maximum number of steps that you want to perform to produce the report?

11. What days of the week would you require the data warehousing system to be available to you?

12. If a particularly complicated ad hoc report required 30 minutes to run,

would you use the report again? If it required 1 hour?

13. What best describes your current use of documentation for software applications.

_____ I frequently use hardcopy manuals

_____ I normally use online documentation only

_____ I will call someone for help using the application

_____ I do not use software applications that are not readily usable

14. Considering the reports you currently receive from the SIS system, how do you go about reconciling questionable results from these reports?

15. How do you prefer to view report data, or what format is most meaningful to you? (Rank in order)

_____ Graphical . graphs, pie charts, etc.

_____ Tabular format

_____ Formatted report

_____ Other (Please explain) _____

16. Considering an analytical processing system, if you had to choose between a system that is available to you during normal working hours and one that is available evenings and weekends, which would be most useful to you?

17. How frequently will you require data to be refreshed in the data warehouse?

_____ Daily

_____ Weekly

_____ Monthly

_____ At the end of each semester

_____ Variable throughout semester

_____ Other (please explain) _____

APPENDIX B

DATA WAREHOUSE EVALUATION

For each of the following questions, please tick the answer that best represents your opinion.

1. There is at least as much information in the data warehouse reports as in reports from the SIS system.

- ☐ a. Strongly agree
- ☐ b. Agree
- ☐ c. Disagree
- ☐ d. Strongly disagree
- ☐ e. Don't know

2. The data warehouse reports contain information that has not been available to me in SIS.

- ☐ a. Strongly agree
- ☐ b. Agree
- ☐ c. Disagree
- ☐ d. Strongly disagree
- ☐ e. Don't know

3. How would you compare the data warehouse reports to SIS reports in terms of understandability?

- ☐ a. Data warehouse reports are harder to understand
- ☐ b. Data warehouse reports are easier to understand
- ☐ c. They are about the same
- ☐ d. No opinion

In data warehouse terms, drilling down is the ability to view summarized data in more details. The Grade Distribution report is an example of drilling down. Please answer the following questions about the Grade Distribution report.

4. The information in the drill down reports is meaningful to me.

- ☐ a. Strongly agree
- ☐ b. Agree
- ☐ c. Disagree

___ d. Strongly disagree

___ e. Don't know

5. The detail information is easier to get to than in SIS.

___ a. Strongly agree

___ b. Agree

___ c. Disagree

___ d. Strongly disagree

___ e. Don't know

6. The drill down reports are harder to understand.

___ a. Strongly agree

___ b. Agree

___ c. Disagree

___ d. Strongly disagree

___ e. Don't know

7 Would you continue to use the data warehouse for reporting?

8. Where possible, would you use the data warehouse instead of SIS for reporting?

9. Would you like to see the data warehouse expanded to include more types of data?

10. In your opinion, what was the best feature of the data warehouse reports?

APPENDIX C

SCRIPT TO DETERMINE THE RECORDS OF THE TABLE

The script below selects the given table and displays the records of that table in ascending order based on the row Record ID.

```
<?php function select()
{
    global $a;
    global $showrecs;
    global $page;

    $res = sql_select();
    $count = sql_getrecordcount();
    if ($count % $showrecs != 0) {
        $pagecount = intval($count / $showrecs) + 1;
    }
    else {
        $pagecount = intval($count / $showrecs);
    }
    $startrec = $showrecs * ($page - 1);
    if ($startrec < $count) {mysql_data_seek($res, $startrec);}
    $reccount = min($showrecs * $page, $count);
?>

<table class="bd" border="0" cellspacing="1" cellpadding="4">
<tr><td>Table: assets</td></tr>

<tr><td>Records shown <?php echo $startrec + 1 ?> - <?php echo $reccount ?> of <?php echo
$count ?></td></tr>

</table>
```

```

<hr size="1" noshade>
<?php showpagenav($page, $pagecount); ?>
<br>
<table class="tbl" border="0" cellspacing="1" cellpadding="5"width="100%">
<tr>
<td class="hr">&nbsp;</td>
<td class="hr">&nbsp;</td>
<td class="hr">&nbsp;</td>
<td class="hr"><?php echo "asset_id" ?></td>
<td class="hr"><?php echo "asset_class" ?></td>
<td class="hr"><?php echo "department" ?></td>
<td class="hr"><?php echo "asset_description" ?></td>
</tr>
<?php
for ($i = $startrec; $i < $reccount; $i++)
{
    $row = mysql_fetch_assoc($res);
    $style = "dr";
    if ($i % 2 != 0) {
        $style = "sr";
    }
}
?>

```

SCRIPT TO DETERMINE AUTHENTIC USER LOGIN

The script below determines whether a visitor to the software is authentically registered in the system as a known user. In the even that the user is not registered it will direct the user to insert his/her user name and password.

The pages also include level based user authentication and therefore limit some aspects of the software to only a certain group of users. i.e. The Administrator Group or the Clerk/Data Entry Personnel. This is displayed in the figure below (after the *code*);

```
<?php function login()
{
    global $_POST;
    global $_SESSION;

    global $_GET;
    if (isset($_GET["a"]) && ($_GET["a"] == 'logout')) $_SESSION["logged_in"] = false;
    if (!isset($_SESSION["logged_in"])) $_SESSION["logged_in"] = false;
    if (!$_SESSION["logged_in"]) {
        $login = "";
        $password = "";
        if (isset($_POST["login"])) $login = @$_POST["login"];
        if (isset($_POST["password"])) $password = @$_POST["password"];

        if (($login != "") && ($password != "")) {
            $conn = mysql_connect("localhost", "root", "");
            mysql_select_db("sis");
            $sql = "select `password` from `users` where `name` = '" . $login . "'";
            $res = mysql_query($sql, $conn) or die(mysql_error());
            $row = mysql_fetch_assoc($res) or $row = array(0 => "");
            if (isset($row)) reset($row);
            if (isset($password) && ($password == trim(current($row)))) {
                $_SESSION["logged_in"] = true;
            }
        }
        else {
            ?>
```


SCRIPTS FOR INSERTING DATA INTO THE DATABASE

```
<?php } ?>
```

```
<?php function showpagenav($page, $pagecount)
```

```
{
```

```
?>
```

```
<table class="bd" border="0" cellspacing="1" cellpadding="4">
```

```
<tr>
```

```
<td><a href="students.php?a=add">Add Record</a>&nbsp;</td>
```

```
<?php if ($page > 1) { ?>
```

```
<td><a href="students.php?page=<?php echo $page - 1
```

```
?>">&lt;&lt;&nbsp;<a>Prev</a>&nbsp;</td>
```

```
<?php } ?>
```

```
<?php
```

```
global $pagerange;
```

```
if ($pagecount > 1) {
```

```
if ($pagecount % $pagerange != 0) {
```

```
    $rangecount = intval($pagecount / $pagerange) + 1;
```

```
}
```

```
else {
```

```
    $rangecount = intval($pagecount / $pagerange);
```

```
}
```

```
for ($i = 1; $i < $rangecount + 1; $i++) {
```

```
    $startpage = (($i - 1) * $pagerange) + 1;
```

```
    $count = min($i * $pagerange, $pagecount);
```

```
if (((($page >= $startpage) && ($page <= ($i * $pagerange)))) {
```

```
    for ($j = $startpage; $j < $count + 1; $j++) {
```

```

        if ($j == $page) {
?>
        <td><b><?php echo $j ?></b></td>
        <?php } else { ?>
        <td><a href="students.php?page=<?php echo $j ?>"><?php echo $j ?></a></td>
        <?php } } } else { ?>
        <td><a href="students.php?page=<?php echo $startpage ?>"><?php echo $startpage ."..."
        . $count ?></a></td>
        <?php } } } ?>
        <?php if ($page < $pagecount) { ?>
        <td>&nbsp;<a href="students.php?page=<?php echo $page + 1
        ?>">Next&nbsp;&gt;&gt;</a>&nbsp;</td>
        <?php } ?>
    </tr>
</table>
<?php } ?>

```

```

<?php function showreclnav($a, $recid, $count)
{
?>
<table class="bd" border="0" cellspacing="1" cellpadding="4">
<tr>
<td><a href="students.php">Index Page</a></td>
<?php if ($recid > 0) { ?>
<td><a href="students.php?a=<?php echo $a ?>&recid=<?php echo $recid - 1 ?>">Prior
Record</a></td>
<?php } if ($recid < $count - 1) { ?>
<td><a href="students.php?a=<?php echo $a ?>&recid=<?php echo $recid + 1 ?>">Next
Record</a></td>
<?php } ?>
</tr>

```