NAVAL
POSTGRADUATE
SCHOOL

MONTEREY, CALIFORNIA

THESIS

DATA WAREHOUSING AT THE MARINE CORPS INSTITUTE

by

Andrew J. Vuillemot

September 2003

Thesis Advisor: Thomas J. Housel
Second Reader: Glenn R. Cook

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2003	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Data Warehousing at the Marine Corps Institute			**5. FUNDING NUMBERS**
6. AUTHOR Andrew J. Vuillemot, Maj USMC			
7. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME AND ADDRESS Marine Corps Institute, Marine Barracks, Washington, DC			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT **Approved for public release; distribution is unlimited**			**12b. DISTRIBUTION CODE**

13. ABSTRACT *(maximum 200 words)*

This thesis is a case study on the value added of an implementation of a data warehouse at the Marine Corps Institute (MCI). Data Warehousing at an environment such as MCI can solve a myriad of strategic questions. The database which MCI possesses, the Marine Corps Institute Automated Information System (MCIAIS), contains a staggering amount of data, waiting to be mined and turned into knowledge for high-level decision makers.

The actual value of the data warehouse to the organization is evaluated using the Knowledge Value Added (KVA) methodology. Many methodologies exist that attempt to measure the value added due to Information Technology. KVA allows description of all process outputs, including those generated from IT in common units. This allows allocation of revenue to IT in proportion to contributions to process outputs at the sub-corporate level, which MCI is at.

This thesis looks at warehouses, ways of measuring value of IT, MCI's organization and core functionality, its current data environment, the implementation of the warehouse, and the value that is added through that implementation.

14. SUBJECT TERMS: Knowledge Value Added (KVA), Data Warehouse, Return on IT, Data Mining, Database.			**15. NUMBER OF PAGES** 103
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

DATA WAREHOUSING AT THE MARINE CORPS INSTITUTE

Andrew J. Vuillemot
Major, United States Marine Corps
B.S., United States Naval Academy, 1987

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
September 2003

Author: Andrew J. Vuillemot

Approved by: Thomas J. Housel
Thesis Advisor

Glenn R. Cook
Second Reader/Co-Advisor

Dan C. Boger
Chairman, Department of Information Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis is a case study on the value added of an implementation of a data warehouse at the Marine Corps Institute (MCI). Data Warehousing at an environment such as MCI can solve a myriad of strategic questions. The database which MCI possesses, the Marine Corps Institute Automated Information System (MCIAIS), contains a staggering amount of data, waiting to be mined and turned into knowledge for high-level decision makers.

The actual value of the data warehouse to the organization is evaluated using the Knowledge Value Added (KVA) methodology. Many methodologies exist that attempt to measure the value added due to Information Technology. KVA allows description of all process outputs, including those generated from IT, in common units. This allows allocation of revenue to IT in proportion to contributions to process outputs at the sub-corporate level, which MCI is at.

This thesis looks at warehouses, ways of measuring value of IT, MCI's organization and core functionality, its current data environment, the implementation of the warehouse, and the value that is added through that implementation.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

The Marine Corps Institute in its entirety must be thanked for the wonderful support that was extended to make this study possible. In particular, Major Scott Sizemore, the head of the Information Technology Department, must be thanked for his initial idea to implement the data warehouse, and the great relationship that was maintained during this project. Additionally, the support of the Executive Director, Mr. Terry Franus, and Ms. Riia Hanson in the IT department was greatly appreciated.

Dr. Thomas Housel must also be thanked for his guidance and support during this project. His trailblazing work in developing the KVA methodology to measure return on IT made this study possible. Glenn Cook must also be thanked for his excellent instruction in the area of economic value of IT, and his guidance and counsel during the thesis process.

Lastly, the faculty and staff of the Naval Postgraduate School must be thanked. This thesis is the culmination of over two years of study at NPS. Given that the thesis covers not only data warehousing, but also measures the value of the implementation, this study drew from almost every course that is available to the IST student. To round out the program drawing on this multitude of subjects made the contributions of all at NPS truly valuable.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

This thesis is a case study of one particular implementation of data warehousing at a military distance learning enterprise, the Marine Corps Institute (MCI), and what value is to be gained from that implementation.

The critical research question is whether data warehouses actually provide any value to the organization. This study will look at data warehouses, what their capabilities and limitations are, and in a specific military proof-of-concept example, what value-added the implementation actually was. The benefit will be measured using the technique of Knowledge Value Added (KVA). This methodology allows descriptions of all process outputs, including those generated form IT, in common units. This allows allocation of revenue to IT in proportion to contributions to process outputs at the sub-corporate level, which MCI is at.

After asking the research question, this thesis looks at data warehousing in general, and compares methods of measuring return on IT. After that, an in-depth look at the problem space, MCI, is done, focusing on the organization and the core processes within that organization.

Following that, MCI's current data environment is reviewed looking at the large target data source for the warehouse, the Marine Corps Institute Automated Information System, MCIAIS. The actual implementation of the warehouse is then chronicled, followed by some of the early returns and projected value to be gained. This data is then analyzed using the KVA methodology to get a true picture of the return on the IT investment in the warehouse by MCI.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION AND RESEARCH QUESTION

A. INTRODUCTION

In the fast moving world of modern business, an entity must adapt to new technologies in order to succeed. Within the last ten years, Information Technology (IT) has grown by leaps and bounds. Computers are everywhere, and with them software applications to do all manner of business operation, from administering websites, to maintaining databases, to executing simple word processing, the desktop of today's business professional seems to do it all.

One would think that with all of this computing power comes a huge increase in productivity. Unfortunately, that is not quite the case. In many cases, the selection, purchasing, installing, and using of IT systems is not worth the cost, both in time and money as compared to what is actually gained in real output. For IT projects, fully 90% have a late delivery date, 75% are over budget, and half never are implemented at all. The U. S. Department of Defense (DoD) alone spent in excess of$50 Billion on IT projects in Fiscal Year (FY) 2000. Added to this is the so-called "Productivity Paradox", which states that huge investments in IT have had little effect on overall productivity. Given all of this grim news, we continue to spend at an ever increasing rate, hoping not only that technology will not pass us by, but also that we will receive some kind of a return on our investment. Although the statistics cited above look bad, many IT projects do add value.

One application that has at times proved to be valuable for enterprises seeking greater productivity is the use of the data warehouse. The data warehouse for business applications exists for several reasons. It allows for data to be extracted from many sources, by "cleaned", and stored into one large data facility. Once in the warehouse, then this data can be arranged and analyzed, and used as a decision support tool to allow strategic decision makers to make more informed and well thought out choices concerning the future of their enterprise.

Many firms have initiated data warehouses for their business. The motivations vary with the organization. Some do it to have the ability to mine their data. Others implement warehouses to have access to the analytical capabilities inherent in the software. Still others only desire all of their "clean" data to reside in one central location. Whatever the rationale, warehousing of data is a popular business option in today's tech savvy world, but is it worth the investment?

B. RESEARCH QUESTION

The real question is whether data warehouses actually provide any value to the organization. This study will look at data warehouses, what their capabilities and limitations are, and in a specific military proof-of-concept example, what value-added the implementation actually was.

C. THE METHODOLOGY

In this study, benefit will be measured using the technique of Knowledge Value Added (KVA). This methodology will demonstrate what the value of the knowledge embedded in the warehouse actually is.

D. THE PROBLEM SPACE

The Marine Corps Institute (MCI) is the distance-learning arm of Marine Corps education. The purpose of MCI is to develop, maintain, administer, track, and ship distance-learning materials to students worldwide. Currently, MCI has over 200,000 students enrolled in active courses. Of these only 58% are active duty Marines. The remainder is made up of Marine reservists, other service members, civilians, retired Marines, midshipmen, and foreign-service members. Fulfilling this mission with only several hundred Marines and Department of Defense (DoD) civilians is a daunting task to say the least. To help manage the task, MCI has harnessed the power of information technology. IT has played a large role in improving overall customer service and streamlining MCI's business processes.

The decision makers at MCI made a decision in 2002 to attempt to use data warehousing to solve some of their issues. First, it would allow them to take their customer database, the Marine Corps Institute Automated Information System (MCIAIS) offline to perform queries and reporting. Second, MCI could use the warehouse's On Line Analytical Processing (OLAP) and statistical analysis tools for support in decision-making. These decisions range from the employment of strategic assets, to more mundane decisions such as proper stocking and delivery of course materials. Third, MCI would use the data warehouse in a data-mining mode to try to unlock the mysteries of the accumulated data in MCIAIS.

Phase I of this process is currently nearing completion. It consists of setting up the hardware and software of the warehouse, and transitioning the overarching MCI database, the Marine Corps Institute Automated Information System (MCIAIS) to the warehouse hardware. Once these tasks are completed, the first implementation tasks will to be to duplicate the existing management reports, which are currently pulled from MCIAIS live. These reports will be published on the Marine Barracks Washington intranet as they currently are published from the live version of MCIAIS. The functionality of being able to pull data from an off-line Warehouse is highly desirable as MCIAIS is a very large, active database measuring over 50 gigabytes in size. Performing large, invasive queries of the data places an inordinate stress on the live database, which is constantly serving live clients. In order to better serve the customer, and allow for a custom arrangement of the data, the concept of the warehouse at MCI was born. A study of the set up of the warehouse will be included in this document in Chapter VI.

Phase II of this process is to determine what potential knowledge can be garnered from the data that already exists. Key individuals at MCI, or the so called "knowledge workers" will be educated, and try to brainstorm possible data relationships that can streamline their operations and add efficiency. Included in this phase is the opportunity to include new data elements that could possibly alter strategic decision making considerably, and make an enormous positive benefit to the Institute as a whole. This topic will be examined in Chapter VI, and a proof of the value of this will be studied further in Chapter VII.

For background material, a brief description of data warehouses and their capabilities will be discussed, as well as a review of value measurement methodologies.

Following that, the subject organization, its critical processes, and the structure of its current database will be examined. Next, the implementation of the warehouse will be chronicled, and the results discussed. Finally, an analysis of how valuable the warehouse actually proved to be will provided.

II. TECHNICAL REVIEW

A. GENERAL

The technical review section of this thesis will look at several key items critical to the research question. First, a brief overview of the data warehouse will be covered, as it is the IT system that's value will be measured in this study. Second, the methods of measuring return of IT will be discussed to illustrate the methodology of how the project will be measured. After both of these topics are looked at, they will be combined to see how value can be measured in the implementation of a data warehouse.

B. DEFINITION OF DATA WAREHOUSES

A stand-alone database, like the live version of MCIAIS before the warehouse, is the place that data is stored. The data is put there for future reference and to run applications, but other than limited querying and reporting capability; it adds no knowledge to our operation. With a warehouse, we are able to extract the data, and more importantly the information and knowledge out by using the warehouses analytical capabilities.

Definitions of data warehouses vary from author to author, and system-to-system based on the capabilities that each warehouse possesses. All seem to agree that a data warehouse is "a database that collects business information from many sources in the enterprise, covering all aspects of the company's processes, products and customers."[1] (Stair and Reynolds, p. 590). Data warehousing has progressed much further than this, however. Just collecting the data into a giant "super database" is no longer the sole function of the data warehouse. A data warehouse is more properly categorized as a Decision Support System (DSS). A DSS, as its name suggests, it a system that assists managers to come to the proper conclusion concerning complex issues. In an environment that has a large amount of data, and a complex set of processes, decisions

[1] Stair, Ralph M., and Reynolds, George W., Principles of Information Systems, Fifth Edition, Course Technology, 2001.

are not always obvious or apparent. Now that the technology exists to make help with decision making available, DSS's fulfill that role.

It must always be remembered that the DSS is only a tool, and not the decision maker. A user of a DSS must maintain control, and use the system wisely.

1. Warehouses

By recognized convention, a "Data Warehouse" concerns the very large database, of a collection of multiple databases or data sources. The conglomeration of all this data into one repository for the express purpose to analyze the data across the organization is the purpose of the data warehouse. A schematic of a building of a data warehouse is shown below in Figure 1.

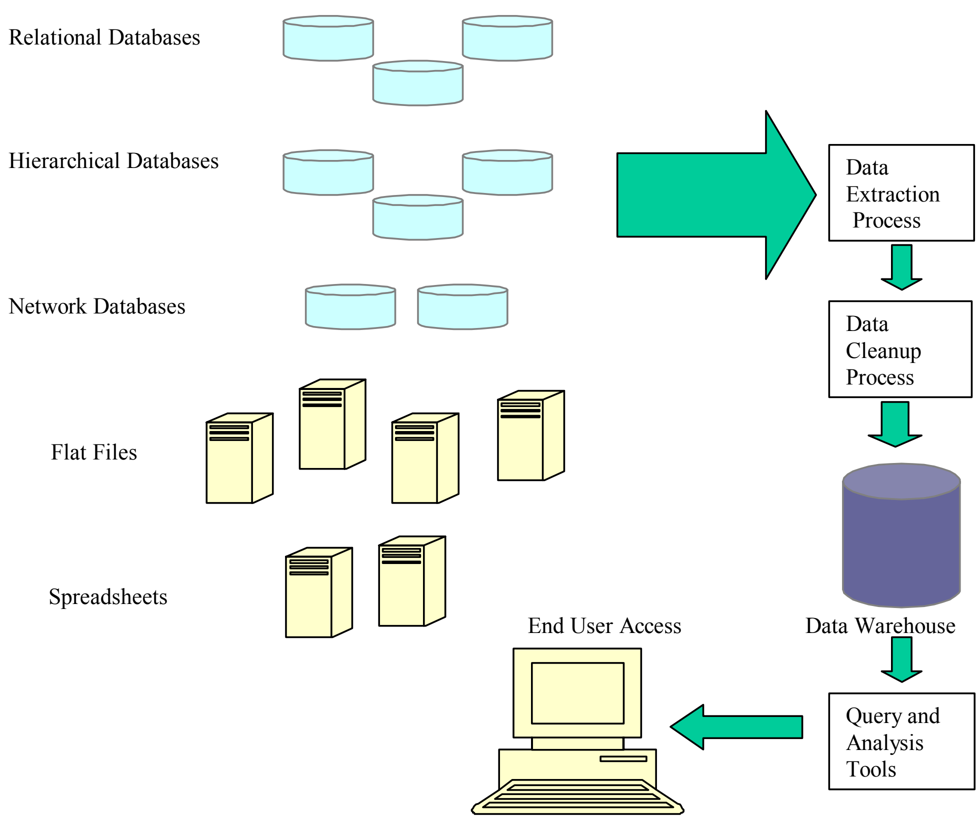


Figure 1. Data Warehouse Schematic[2]

[2] Ibid, Stair and Reynolds, p. 192.

The figure illustrates how the data is extracted form multiple sources such as relational and hierarchical databases, flat files, and spreadsheets. Once extracted, the data is "cleaned" by eliminating any old, irrelevant, or duplicate data from the combined set. At that point, it is stacked into the data warehouse to be manipulated by the querying and analysis toolkit that comes with the warehouse software. Those products are then distributed to users, usually via the three-tier architecture.

2. Data Marts

As opposed to the warehouse, a data mart is either a subset of the warehouse as a whole, or a smaller version of a warehouse. If a subset, the data mart is one particular department of the organization as a whole's data, containing the capability to analyze that data with the context of the business processes within that department. Figure 2 below shows an example of this configuration of a data warehouse containing data marts.

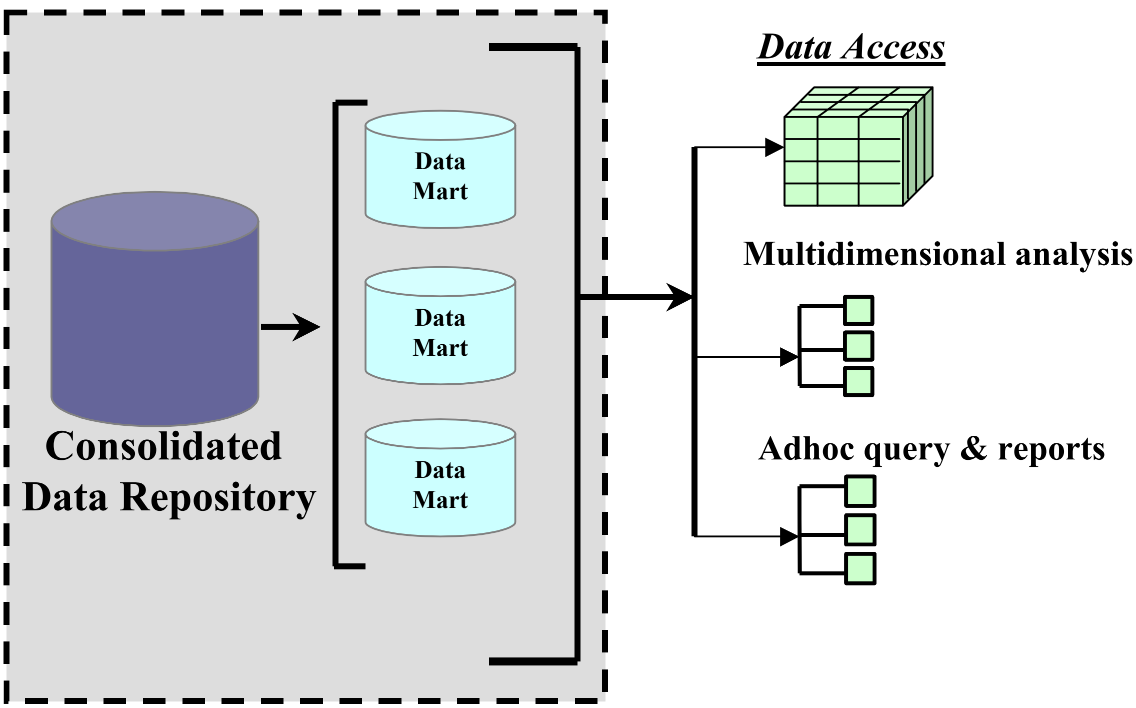


Figure 2. Data Mart Example[3]

[3] medicine.osu.edu/Informatics/talks/ powerpoint/CCG%20Presentation.ppt, August 2003.

In the MCI example, data marts could exist for logistics, course development, and customer service, each being a department of the whole, and not needing the entire database to achieve the desired functionality. A standard commercial example would have a large overarching data warehouse, with data marts for accounting, human resources, shipping, manufacturing, etc.

Some smaller enterprises, such as “Mom and Pop” stores, who have data less than 20 gigabytes or so, may build a scaled down version of a warehouse to suit their needs. Often times, these “mini-warehouses” are referred to as data marts. By using the term data mart in this case, it is not implied that any functionality is lost. The only difference in a case such as this is that the size of the warehouse is smaller, so the term data mart is used in place of data warehouse.

C. FUNCTIONS OF A DATA WAREHOUSE

In addition to the storage capability of all aspects of a company’s data, as described in the definition, there are retrieval and analytical functions which exist as well. They include querying and reporting, On-Line Analytical Processing (OLAP) and statistical analysis, and data mining.

1. Queries and Reports

The most simple of the data warehouse functions are querying and reporting. This is simply asking the database for specific information from the data fields, and displaying it. In the MCI example, it could be something as simple as asking MCIAIS how many courses a specific individual has enrolled in, how many he has completed, and the titles of those courses. The queries can grow more complex, but they never evaluate the data, they just report back on what is asked for.

a. Queries

A query, like it sounds, is a question. All database programs, however basic have a query function built in. A simple query is to ask a database for a list of Marines who are 30 years old. This will work, as long as there is a table of Marines, and age is one of the attributes listed in that table. The answer will be the list of Marines, all of whom are 30. More complex queries ask questions across different tables in the database. One possible query of this nature would be how many Marines are 30, and are enrolled in the MCI course 08.20. The query, either asked directly using the functionality of the database, or using Structured Query Language (SQL), would ask the database to search the table Marine for those with age 30 and the table Courses Enrolled for those specific Marines. The link between these two fields would be some descriptive element about the Marine, usually a number identifier like a student number or a social security number that exists in both tables. This logic carries on for more complex queries across numerous tables and specific fields.

From a management perspective, it is rather obvious that to know grouped data elements, such as how many Marines are 30 can be useful. Drilling down for more specific data in more advanced queries can alter strategic decisions at the upper levels of the organization.

b. Reports

Reports are nothing more than the presentation of the query that is asked. Usually, the management user of the query has a very limited knowledge of how a database operates. His concern is the decision he is going to make concerning how many 30-year-old Marines are enrolled in course 08.20. He doesn't want to see how the information technology person manipulates the database, and he surely doesn't want to see the query written in SQL. He just "wants the facts". With a report, a nicely laid out version of the results of the query is presented for the individual's use.

With the introduction of many company intranets, many commonly used management reports are provided in an online version for all users to see. This allows the IT section to develop the query and the report one time, and have it posted in a location that all can see. This functionality exists in any concern that has a database on an intranet server. The database is queried, the results are placed into a formatted report, and that report is made available to those who have access to the intranet.

2. OLAP and Statistical Analysis

To look across data multi-dimensionally, OLAP alone, or combined with statistical analysis can be used. Performing this processing allows the user to drill down deeper into the data sets, and pull out any underlying relationships, and to what degree these relationships apply. Normal databases do not possess inherent OLAP capabilities. To get analysis of the data, and to arrange the data to see across multiple dimensions, some kind of additional software functionality is required.

a. OLAP

On Line Analytical Processing is essentially a high-speed way to look deep into multidimensional databases. Querying and reporting are good at looking at relatively simple relational databases, across a limited scope of the organization. When the database becomes very complex, has multiple layers, covers an entire operation, and is of a very large size, writing the SQL query to know how may widgets were produced in Topeka last quarter, that were sold in Fairbanks, shipped via FedEx the previous quarter at the bargain price set would prove to be a very difficult task. OLAP provides the user with the tools to perform this “drilling down” at high speed.

The reason OLAP works in data warehouses is that the warehouse is set up in order to maximize the OLAP capability. The organization keeps OLAP in mind when entering the data into the warehouse form the data sources. In essence, the data is “stacked” with the idea of analysis in mind, much like a Marine Corps unit loads a ship for amphibious assault. Marines pack units into amphibious cargo and air-transportable spaces, with the mission in mind. Units that are to be sent in first are loaded last. Flexibility is allowed by modularizing the load plan to give the commander different landing options with a minimum of movement of the load. Such is true when the warehouse is actually loaded. Key data elements that affect other business processes are loaded in so to best make use of the OLAP analytical capability. Because of this, stacking and the speed now available, business questions that would take multiple queries and extensive digging are available in an instant.

b. Statistical Analysis

Statistical analysis is nothing new. Anyone who sat through a college statistics class knows all about normal and Poisson distributions, and how to look at data from a statistical perspective. Doing this analysis is often a very painful exercise, as the mathematics involved is complicated and difficult for the average business user. The technical revolution put many statistical tools into our hands, such as Microsoft Excel. Unfortunately, most Excel users can only scratch the surface of the programs capabilities, and have a difficult time doing that. When you factor in trying to take the output of a multidimensional query and performing that analysis, it becomes almost impossible.

Many data warehouses have built-in capability to perform the analysis on these difficult data sets. This capability, in conjunction with OLAP, makes the data warehouse's controller powerful indeed. He can now pull information out of a large data source across the entire organization, see how multiple dimensions affect it, and analyze the outcome in a very short amount of time.

3. Data Mining

Data mining is utilizing a set of tools that allow the user to discover hidden patterns that exist in the data that are not obvious to the casual observer. As it's name states, the database is a quarry of information, and the miner is the user. Nuggets of gold are there, if the user chooses to mine them. This process does not require the user to analyze the data structure, and ask specific questions of the warehouse. Instead, data mining uses built in analysis tools that generate and test hypotheses concerning the data set. In layman's terms, the mining tool tries to find the gold where you haven't thought to look.

A classic example of data mining is the use of the grocery store frequent shopper cards at the supermarket. You buy your groceries, and scan the card to get a discount. The computer scans the UPC codes of the products, and you purchase is recorded in the database. The data mining tools then go to work, trying to find connections that the normal manager could never put into a logical query. The famous example is where men most buy beer and diapers together. These two products would seem to be unrelated, in fact most advertisers and store managers would never put these items in the same category. A data mining tool proved that there was indeed a correlation between the two, and smart supermarkets then market the two together, or give the customers coupons for both, anything to entice the buyer to choose their supermarket over another one.

D. COMMON USES OF A DATA WAREHOUSE

There are several categories of entities that will turn to warehousing to fulfill their needs. They include industries that implement Customer Relationship Management (CRM), high-trust, high-value businesses such as banks and insurance companies, and other concerns that have a large amount of data that can be analyzed by the warehouse, such as distance learning institutions.

1. Customer Relationship Management

A common use in industry is with Customer Relationship Management. What CRM tries to accomplish is the better treatment and hopeful retention of customers by better catering to their needs. By using the data concerning those customers to encourage and promote future sales, CRM has become big business. The sheer magnitude of data collected in CRM leads to a huge database. To effectively extract all of the combinations of queries, as well as to mine for new ways to please the customer, and thus increase profits, data warehousing is implemented.

A good example of a company that uses CRM is the Saturn division of General Motors. They promote themselves as the small company that is different. They manufacture their cars with the customer in mind, and do little things like sending you a birthday card for you car when it turns a year old. There no-haggle pricing system allows buyers who hate to deal with salesmen to walk in to a showroom and remain unmolested. Some of these ideas were the brainchild of a marketing guru, however Saturn maintains its edge in dealing with its customers using data warehouse CRM technology.

Given that it cost a huge amount to attract a customer, and far less to retain one that has already purchased an item, CRM is a strategy that has been proven to save money and earn more revenue. The cost of the CRM solutions are almost always made back in a very short amount of time.

2. Other Warehouse Candidates

Banks and insurance companies also rely on data warehousing, to sift through the mountains of data and identify trends in their customer base. This is used to determine risk, detect fraud, and keep the cost of doing business low while still maintaining optimal operating conditions.

Distance learning educational entities are also perfect candidates for data warehouses. The volume of information concerning development and production of materials, as well as data concerning student enrolment and course completion can easily be harnessed to make smarter strategic decisions.

E. COMMON PROBLEMS WITH DATA WAREHOUSES

Many entities with large sets of data feel that by buying a data warehouse that all of their problems will be solved, the warehouse will pay for itself in a matter of months, and they will have single-handedly saved their enterprise. Like most IT solutions, data warehouses are a tool, that if used at the right place, at the right time, may bring about improvements to operations. Many cases do not. What follows is some of the common pitfalls associated with data warehouses.

1. No Need for a Warehouse

The first and most common error is that a data warehouse solution just doesn't fit the enterprise that is installing it. Whether too big, or too small, size does matter. A large company with mountains of data in many disparate forms and functions to huge to count is probably too big for a warehouse. This may be data warehouse sacrilege, but other, more appropriate solutions such as an Enterprise Resource Planning (ERP) tool is more likely the appropriate response. For those enterprises that contain all of their data is one or several small databases, and don't have many departments or functions, a data warehouse may just prove to be a very neat toy that is only utilized at a fraction of its capability. A true data warehouse candidate needs to be a robust organization with sizeable data and many functions to truly pay off in the end.

2. Source Data is Unavailable

IT professionals in this scenario promise the stars but are unable to deliver because the data they need for implementation is just not there. They may think that they have all that they need in their legacy systems, but when the time comes to extract, transfer, and load, they cannot produce all of the data they thought that they had. This lack of a historical base often cripples the warehouse on the launching pad.

3. Lack of User Buy-In

This fatal flaw kills many IT solutions from databases to ERPs. If the user doesn't understand the benefits of the warehouse, and fails to use its products, or is so tied to the "old way" of doing it, the benefits will never be achieved. With any change in business processes, buy-in from the top it critical to success. Management needs to let all hands know that they support the warehouse fully, and expect to see the common user to produce results with this new technology. If this doesn't happen, the time, money, and effort expended in the warehouse is all wasted and an expensive pile of junk is back in the server room.

4. Poor Implementation

Implementers that don't know what they want to ultimately achieve, or don't really know what they are doing can be a black hole that resources are poured into with no return. An entity may be a perfect warehouse candidate, but if they don't know what they want, the setup will not achieve results anywhere approaching maximum. Also, it the team that implements the solution is technically lacking, the setup will never get off of the ground. A picture of monkeys trying to play football nicely illustrates this point. Additionally, the warehouse cannot be seen as an IT only project. Often times, IT doesn't understand fully what the operators are doing. If the workers who perform the critical core processes are not sufficiently involved, IT may get a good tool for its own use on IT specific areas of the entity, but the main value of the warehouse will be squandered.

F. MEASURING THE RETURN ON IT INVESTMENT

As stated in the introduction to this thesis, just by investing in an IT solution such as a data warehouse does not necessarily mean that a huge gain in productivity is going to logically follow. Millions upon millions of dollars have been invested by corporate entities worldwide on IT that has never produced any results. In fact, with many systems implementations, the time and money spent to "upgrade" the IT infrastructure has actually lost money, and decreased productivity.

Be it cubit, yard, or meter, distance is measurable. It is also the case with the return on IT. There are many ways in which corporate entities measure how well they have done with their investment. One traditional method is the accountant's standard: Return On Investment (ROI). Although calculated in a more comprehensive way in reality, the basic model is very simple. A corporation spends X. Over time, the additional profit gained from implementing the system is Y. The total of Y minus X equals the ROI. But is it that simple? If it were, there wouldn't be an industry that has grown up around measuring how much return organizations are getting from their IT systems.

Just like the example above using cubits, yards, or meters, there are different ways to look at return on IT investment. The question is which method of measurement gives the most accurate results? Table 1 below lists several approaches to measuring value of return on investment on IT.

Approach	Focus	Example	Level of Analysis	Key Assumption	Key Advantage	Limitation
Process Of Elimination	Treats effect of it on roi as a residual after accounting for other capital investments	Knowledge capital (Strassmann 2000a, b)	Aggregate corporate - level only	ROI on it difficult to measure directly	Uses commonly accepted financial analysis techniques and existing accounting data	Cannot drill down to effects of specific IT initiatives
Production Theory	Determines the effects of IT through input output analysis using regression modeling techniques	Brynjolfsson & Hitt (1996)	Aggregate corporate - level only	Economic production function links IT investment input to productivity output	Uses econometric analysis on large data sets to show contributions of IT at firm level	"black-box" approach with no intermediate mapping of IT's contributions to outputs
Resource-Based View	Linking firm core capabilities with competitiveness	Jarvenpaa & Leidner (1998)	Aggregate corporate - level only	Uniqueness of IT resource = competitive advantage	Strategic advantage approach to IT impacts	Causal mapping between IT investment and firm competitive advantage difficult to establish
Option Pricing Model	Determines the best point at which to exercise an option to invest in IT	Benaroch & Kauffman (1999)	Corporate/ sub-corporate	Timing exercise option = value	Predicting the future value of an IT investment	No surrogate for revenue at sub corporate level
Family of Measures	Measure multiple indicators to derive unique contributions of IT at sub-corporate level	Balanced score-card (Kaplan And Norton, 1996)	Sub-corporate	Need multiple indicators to measure performance	Captures complexity of corporate performance	No common unit of analysis/ theoretical framework
Cost-Based	Use cost to determine value of information technology	Activity-based costing Johnson & Kaplan (1987)	Sub-corporate	Derivations of cost ? value	Captures accurate cost of IT	No surrogate for revenue at sub corporate level -- no ratio analysis
KVA	Allocating revenue to IT in proportion to contributions to process outputs	Housel & Kanevksy (1995)	Sub-corporate	IT contributions to output ? IT value-added	Allocates revenue and cost of IT allowing ratio analysis of IT value-added	Does not apply directly to highly creative processes

Table 1 Common Approaches That Can Be Used To Measure The Return On IT[4]

Several of these methods look only at the corporate level, while others look at the much lower sub-corporate level, and some have the breadth to look at both. What follows is a description of each method; some of its strengths and weaknesses, and on what level it measures the value of IT.

1. Process of Elimination

The Process of Elimination method is looking at the organization's overall ROI, and after accounting for any other measurable factors; any remaining return must be

[4] Housel, Thomas, J., et.al., "Requirements for a Theory of Measuring the Return on Information Technology: A Proof of Concept Demonstration, December, 2002, p.6.

attributed to return on IT. This makes sense from the accountant's point of view, but since ROI analyses often miss critical factors, and since return on IT is so hard to quantify, this methodology does not produce the most precise results in most cases. This method is at the corporate level only.

2. Production Theory

This methodology looks at the big picture at the corporate level only. The major assumption at work is that IT in equals some productivity out. This works fine if you are the CEO of the corporation and you want to know the approximate results of your IT investment, however if you are at the sub-corporate level and want to know just how much specific IT initiatives are working, this method does not give you the insight that you need.

3. Resource-Based View

Like the Production Theory described above, the Resource-Based View is a corporate level method of looking at return on IT. Its supposition is that by possessing the new IT product or service makes the corporate entity stronger and more competitive relative to rivals. Since competitive advantage is such a murky prospect to establish, especially at lower levels, this competitive advantage is almost impossible to strip out of the equation.

4. Option Pricing Model

Much as this model sounds, it is like exercising stock options at the precisely right moment to maximize profits. The Options Pricing model predicts the future of IT development and chooses the optimal point in time to make the investment to maximize value. This model works at the corporate as well as lower levels, but is difficult to measure the lower that you go.

5. Family of Measures

The Family of Measures approach is one that tries to take into account as many factors as possible to produce the most complete report on the return on IT at the sub-corporate level. It is difficult, however to determine true value as there is no theoretical framework and no common units of analysis. What this means is that although a multitude of factors are looked at, the measurer doesn't know how each statistic relates to the others. This brings forth subjective analysis, and depending on the researcher, can bring vastly differing results.

6. Cost-Based

The framework for the Cost-Based methodology is rooted in ROI, and specifically shown in Activity-Based Costing (ABC). A very thorough and methodical process, Cost-Based methodology attempts to gives the user a true dollar value on every activity, thereby making everything equivalent to dollars. This seems like a good approach, but at times misses the mark on analysis of true value. There is value in many things, and ABC and other Cost-Based systems cannot truly quantify nebulous items such knowledge in fiscal terms. This method is primarily used at the sub-corporate level.

7. Knowledge Value Added

KVA allows description of all process outputs (including those generated from IT) in common units. This allows allocation of revenue to IT in proportion to contributions to process outputs at the sub-corporate level. What this does, is show in measurable terms what value is actually added by implementing specific IT applications. Additionally, KVA allocates revenue as well as cost allowing for ratio analysis of the IT value-added. This methodology is not directly applicable to highly creative processes, but for most IT implementations, it is a very good gauge of what bang for the buck is achieved.

G. DATA WAREHOUSES AND RETURN ON IT INVESTMENT

Having outlined the earlier part of this chapter looking into what data warehouses are and are capable of, and also looking at ways to measure return on IT investment, it is now in order to combine the two and look at the return on data warehouses.

1. The Vital Question

Of the utmost importance is answering the question: "What will the data warehouse actually do for my enterprise, and is the investment worth the expense?" A concise answer to the question is as follows:

> A well-designed reporting solution determines what information is needed to support core business processes. The solution can often result in a data warehouse to take the strain off administration systems and allow for enhanced user and report functionality. The value of data warehouses can be measured in decreased support time, increased reporting capabilities and maximized value of new and existing systems.[5]

But how much is an organization willing to pay in real dollars for things such as decreased support time and increased reporting capabilities? What are these things actually worth? That is really the question.

2. Measuring Value

The dollars that are spent to implement a data warehouse are very easy to measure. Hardware, software, and the labor of the people who install and implement all cost real money that comes from the corporate coffers. That part of the equation is very easy to quantify. In the simple equation of value gained over money spent, the cost being in the denominator. The numerator is a different story. If we went back to simple ROI, money in over money out, then it is easy, but with IT solutions that just is not the case. In the methodologies listed in the previous section, many implicitly assume a numerator, but none provide one at the sub-corporate level.

[5] http://www.ceiss.org/products/case-odu.asp, November, 2002.

3. Proof of Concept

Given the multitude of enterprises, specific IT systems that they might be using, and what each of those entities measure as valuable, a virtual myriad of combinations exists. The organization must ask itself what they hope to achieve, and how much they are willing to invest to reach their goals. This study looks at one specific military example using the KVA methodology to determine what the true value of implementing a data warehouse actually is to that organization.

KVA was selected as the measuring methodology for several reasons. First, KVA measure how much knowledge value is actually added. In an enterprise such as MCI, this measurement is critical to the determination of success. As a sub-corporate enterprise, belonging to the Marine Corps as a whole, MCI is well suited in size for KVA. Additionally, the down side of KVA, not being able to adequately measure highly creative processes is not at play since all aspects of MCI's operations are straight-forward. Additionally, KVA fits the mold of MCI due to MCI being sub-corporate, looking internally at its own data, and retrospective in nature or looking at the history of its data. Table 2 below shows that KVA puts the X in the right block for MCI's purposes.

KVA (example)	Corporate		Sub-Corporate	
	Prospective	Retrospective	Prospective	Retrospective
External				
Internal		X		X

Table 2 Approach Focus[6]

[6] Housel, Thomas, J., et.al., "Requirements for a Theory of Measuring the Return on Information Technology: A Proof of Concept Demonstration, December, 2002, p.16.

III. MARINE CORPS INSTITUTE

A. GENERAL

In order to get an idea of the problem space in which the data warehouse will be implemented, the structure of the Marine Corps Institute must be discussed in order to see how the organization fits with the tasks it must perform.

B. THE MISSION OF MCI

The organization of MCI is unique due to its location. MCI is an integral part of Marine Barracks, Washington, D.C., whose primary mission is to provide ceremonial support to the nation's capital. Due to this relationship, the Commanding Officer of Marine Barracks, Washington is the titular head of MCI. Its day-to-day operations are left to the Deputy Director, MCI. The official mission reads as follows:

> Since February 1920, MCI has facilitated the training and education of individual Marines anywhere, anytime. To that end, MCI ensures access to products and provides opportunities to improve performance, to enhance Professional Military Education, and to provide promotion opportunity, together with sponsors of Marine Corps education and training programs. Concurrently, MCI coordinates and executes the Hosting and Parade Escort plan for the Evening and Sunset Parades, while simultaneously providing ceremonial Officers and SNCOs for the Parade Staffs and other assigned ceremonies in order to promote the Marine Corps' heritage and to enhance the Marine Corps' image to the general public. MCI Company also maintains Individual MOS and Battle Skills proficiency both in garrison and field environments to prepare the individual Marine for combat.

There is a significant emphasis in the mission statement that does not concern development of learning materials, but talks of things such as parades and ceremonies. That is a major part of MCI, and causes unique challenges.

C. THE ORGANIZATION OF MCI

MCI is organized with a headquarters and three departments. The schematic of this organization is shown below as Figure 3.

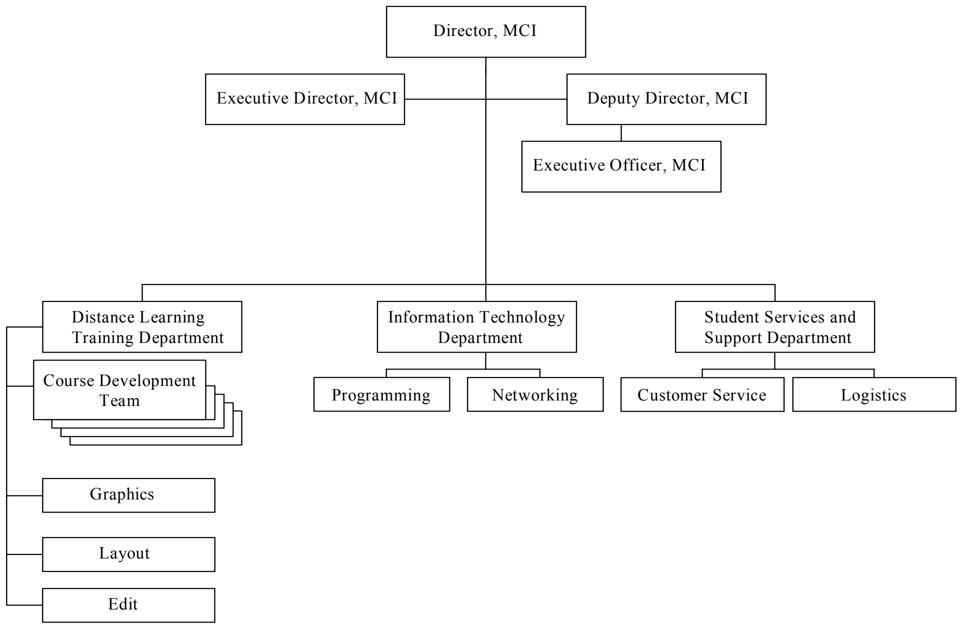


Figure 3. Organization of MCI

1. Headquarters Section

The headquarters of MCI is headed up, as previously stated, by the Commanding Officer, Marine Barracks, Washington, D.C. The running of MCI is the primary responsibilities of the Deputy Director, a Marine Lieutenant Colonel, and the Executive Director, a civilian GS-15. They are assisted in this by the Executive Officer, a Marine Major.

2. Distance Learning Training Department (DLTD)

DLTD is the department that produces course content. It is headed by a GS-14 civilian, who is assisted by a Marine Captain. Included in DLTD are five Course Development Teams. Each is composed of a Project Manager, usually a Marine Captain, an Instructional Systems Specialist, usually a GS-12 or GS-13, and from five to eight Marine Staff Noncommissioned Officer (SNCO) Distance Learning Instructors (DLIs). These teams are organized into Military Occupational Specialty (MOS) groupings that include infantry, combat support, combat service support, administration, and communications. In addition to the teams, separate entities exist within the department for the purpose of assisting the course developer with the integration of graphics, proper layout, and edit.

3. Information Technology Department (ITD)

ITD controls all of the information systems within MCI. It is headed by a Marine Major, who must possess a Master's Degree in Information Technology Management. The department is broken into two divisions, one for programming and one for networking. This department operates and maintains the Marine Corps Institute Automated Information System (MCIAIS) database.

4. Student Services and Support Department (SSD)

This department, headed by a GS-14, is broken into two divisions, one that handles customer service and one that handles logistics. The Customer Service Division is charged with enrollments of courses, examination grading, and interface with the customers via e-mail or telephone. The Logistics Division handles matters such as printing and reprinting of course materials, fiscal matters, postal issues and whatever other logistics matters that the Director sees fit to assign the Division Officer.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. BUSINESS PROCESSES

A. GENERAL

Now that a technical review has been conducted and the organization of MCI has been discussed, an actual description of what MCI does must be chronicled. To understand what the data contained in the warehouse will be used for; the processes under which MCI operates must first be understood.

B. HEADQUARTERS SECTION

Executive As previously stated, the Headquarters Section consists of the Deputy Director, the Director, and other individuals tasked with assisting those two key individuals. Their task, as with any headquarters, is to ensure efficient operation of the organization.

1. Strategic Decision Making

The critical decisions that must be made at MCI are made at the Headquarters. These critical decisions range from investing in new technologies, such as buying a Data Warehouse capability, upgrading software and hardware, determining which courses need priority for development, and proper organization of personnel. These decisions impact how MCI will perform both in the short and long terms. It is critical that the Headquarters is provided with all of the information necessary to make these calls. If a priority is placed on a course that has a very low enrollment, higher enrollment courses, and therefore more students suffer. Likewise, if an expensive IT system is invested in, and it is a system that does not meet the needs of MCI, it is money wasted. It is at this level that MCI has to have the critical data, organized and filtered to properly provide decision support. The data warehouse, with its analysis and mining capability will be able to filter through much of this information to provide the executives with up-to-date, valid information with which to make more informed decisions.

2. Allocation of Assets

Hand-in-hand with decision-making is asset allocation. First and foremost is the allocation of MCI's most precious commodity: people. Like any organization within the military, MCI has a limited amount of personnel with which to perform its mission. Within DLTD especially, a course development team cannot work on five priority projects at the same time. The allocation of work to these teams is critical to the successful functioning of this organization. Additionally, the budget is finite as well. High-ticket items such as computer hardware and software have to be weighed with more mundane items such as administrative supplies and training packages for the Marines and civilians of the Institute. Having the knowledge to make these decisions is crucial to the Headquarters of MCI.

3. Annual Plan Development

A specific example of the MCI headquarters in action is the development of the annual plan. Like all governmental agencies, MCI must deal with the reality of being tied to a budget that works on the fiscal year cycle. Because of this, a new plan must be developed and implemented each year in order to appropriately allocate resources for that period of time. MCI uses a seven-step process to come up with this plan of action. In MCI's current data configuration, much of the analysis conducted in the analysis phases is drawn directly from MCIAIS. The implementation of the data warehouse will streamline these steps considerably and provide for more timely analysis and a wider range of options in the analysis. The comparison of the value added will be examined further in Chapter VII. The steps in Annual Plan Development as they currently exist are outlined below.

a. Strategic Vision Conference

The initiation of the process is a conference held by the Executive and Deputy Directors, and includes the heads of the three departments, and any other key individuals necessary to the process. This is a kick off meeting where analysis tasks are assigned to specific departments, and guidance is issued for the development of a successful plan.

b. Previous Plan Analysis

Like any organization, learning from history is an important step in avoiding previous mistakes. Previous plans are examined for what worked well, and what didn't quite go off as planned. These lessons are taken forward and incorporated into the new plan.

c. New Requirements Analysis

MCI serves a vast customer base as well as many outside agencies such as Training and Education Command, as well as MOS-specific agencies. New initiatives as well as needed revisions are often suggested or in some cases mandated by these customers. These requirements must be factored in to the new plan as they have a critical impact on asset allocation for the coming year.

d. Manpower Analysis

MCI maintains the same Table of Organization strength form year to year with only minor modifications. With this relatively constant pool of manpower and more importantly MOS expertise, they must determine which courses will be worked on and which will remain untouched in the upcoming campaign.

e. Fiscal Analysis

As with any business, budgetary realities exist at MCI. Manpower costs are not nearly the concern they are in the corporate world as military pay is not within the scope of MCI's budget. Operating expenses, as well as logistics costs and new innovations such as IT upgrades must be taken into account. What can be accomplished in the next campaign must fit within the fiscal box that is drawn for them by their higher headquarters.

f. Analysis Conference

Once all of the requirements pieces are assembled, these factors are weighed against each other and at the Analysis Conference differences are hammered out to get a rough plan in place.

g. In-Process Revision

Any questions that arise during the Analysis Conference that require further analysis are then thrown back to the department heads for revision. Once these questions are sufficiently answered, a final Plan Conference is scheduled.

h. Plan Conference

The final Plan Conference is held when all analysis is complete and the headquarters issues its plan of the upcoming campaign.

C. DISTANCE LEARNING TRAINING DEPARTMENT (DLTD)

As stated in the previous chapter, MCI is a distance learning entity. Its existence is based upon providing the students with educational materials in order to educate themselves about professional topics. A wealth of information exists in MCAIS, the current database, concerning many aspects of courses. Which courses are active, how many students are enrolled, failure and dropout rates, who is taking the courses, and all sorts of information concerning the examinations are present. DLTD by properly utilizing and mining this data via the data warehouse can make large strides in optimizing its business practices.

1. Types of Courses

Within MCI, there are three types of course materials: Professional Military Education (PME), Performance Support Tools (PSTs), and Military Occupational Specialty (MOS) Courses. All of these materials are developed and maintained within DLTD at MCI. Course materials have some very important significance to the individual Marine. Completion of any course with an examination can earn junior enlisted Marines points toward promotion. Once the Marine becomes more senior, the carrot is replaced with a stick. Failure to complete specific PME courses can deny a Marine promotion.

a. PME

PME is a very important part of what MCI does. The courses that are contained in this classification are of a professional development nature, and in all cases are tied to the promotion of the individual student as described above. Theses courses include Fundamentals of Marine Corps Leadership, the Marine Noncommissioned Officer, the Career Staff Noncommissioned Officer, the Advanced Staff Noncommissioned Officer, the Warfighting Skills Program, Amphibious Warfare School, and the Command and Staff College.

These courses serve a variety of functions. Fundamentals of Marine Corps Leadership was written for young enlisted Marines. Its education grounds the Marine in what it is to be a leader, and prepares them for promotion to Non Commissioned Officer (NCO).

The courses for NCOs and Staff Noncommissioned Officers (SNCOs) are primer instruction. In order to take less time on classroom instruction and maximize the resident experience, these courses are completed prior to attendance of the resident school of the same name.

The Warfighting Skills program is designed for senior SNCOs and Warrant Officers. This course serves to educate the senior enlisted leaders and long-term experts in warfighting doctrine, and the basic fundamentals of Marine Corps Planning.

Amphibious Warfare School and the Command and Staff College programs exist as alternatives to resident education. Each of these programs is considerable in length, as they mirror one year's worth of resident education.

As was mentioned previously, these courses are tied to a Marine's promotion. If a Marine has not completed his appropriate PME for his grade, he will automatically not be considered fully qualified for promotion and fail selection. Due to this fact, these courses are extremely important to Marines of all grades.

The content of these courses are developed at the Training and Education Command at Quantico, Virginia. MCI personnel assist in editing, educational review, layout and graphics, but no real course content is developed by the Marines of MCI.

b. Performance Support Tools (PSTs)

Performance Support Tools are not specifically designed to educate in the distance-learning model. Instead, these materials are stand-alone job performance enhancers. Due to the expeditionary nature of the Marine Corps, it is not always feasible to carry large volumes of publications to combat areas. What the PST attempts to do is give the Marine a small-sized "cook book", or pocket handbook to help with complex tasks.

A good example of this is the Forward Air Controller handbook. This pocket-sized book, printed on weather-resistant paper, lists all of the aircraft in the U.S. inventory, what armament they are capable of carrying, and instructions on how to properly conduct an air strike. Other useful information such as procedures for other supporting arms such as artillery and naval gunfire is also present. Additionally, useful hints on operating communications equipment and blank pages for notes are also contained.

The content of these PSTs are decided upon by the Content Development team, and validated through Training and Education Command, as well as Marines in the fleet.

c. MOS Courses

Military Occupational Specialty Courses are the bread and butter of DLTD. These traditional paperback course books teach Marines in how to tasks in a specific warfare community, or how to perform the mission with specific equipment.

The list of titles within MCI is strongly tilted toward these types of courses. These courses are significant for Marines in that they allow additional education on one's primary MOS, allow the student to learn about additional equipment types within their specialty, or facilitate study in areas that a Marine has no formal school training, but desires to achieve proficiency. For example, a Marine artilleryman straight out of Artillery School who wants to possess more in-depth knowledge of his gun could enroll in "The M198 Howitzer" course. That same artilleryman, now about to move up to the duties of a Platoon Sergeant, could take "Marine Artillery Survey"; to better learn those functions of his new billet. To be a well-rounded Marine, he could take "Forward Observation"; so if the occasion arose, he could competently call in fire missions, not just fire them.

2. Course Media

Courses from MCI exist in two different types of media. They are in book form, or in Interactive Multimedia Instruction. Internet online instruction is envisioned for the future, but to date, no such courses are available for enrollment.

a. Book Courses

The traditional method of providing distance learning, books that are distributed, studied, and tested on, make up the bulk of MCI's available instruction. These courses are rather easily developed, require no computer for instruction to take place, and are inexpensive as compared with their digital alternative. MCI will continue to develop and maintain this medium of courses far into the forseeable future.

b. Interactive Multimedia Instruction (IMI)

IMI has existed at MCI since the early 1990s. The first course developed was designed for adminstrative clerks, to enhance their performance on their key work tool: their computer. It was as distributed on a 5 ¼" floppy disk to students who obviously had access to a computer for the instruction, as many students at that time had no access. The changes in technology in the 1990s allowed for the development of more titles on better media. CD-ROMS are now produced to give the student a more visual learning experience in a multi-media environment. Course development in this area is still restriced in some part due to computer access. Young Marines do not routinely have access to computational devices capable of running the software, as Marine Corps offices are limited in what computers they are allowed to possess. To offset this somewhat, Distributed Learning Centers (DLCs) have been established at major Marine Corps facilities worldwide to give access to computers to the young Marine. Due to high deployment rates, it is often not possible for Marines, especially those in ground combat units, to take advantage of these opportunities. Because of this MCI has limited its production of IMI courses to those specialties that work in offices and have ready access to computers.

3. Course Development

The core process of DLTD is the development of course materials. This multi-step process ensures that the materials available for enrollment are complete, correct, and up-to-date. The current process uses IT as much as possible given the current IT assets that are available to the users. Significant improvements in the return on IT can be achieved, especially in the completion of the analysis and in the conducting of the conferences by using the capabilities of the data warehouse. An analysis of this improvement will be discussed in Chapter VII. The process of course development is outlined in Figure 4 below.

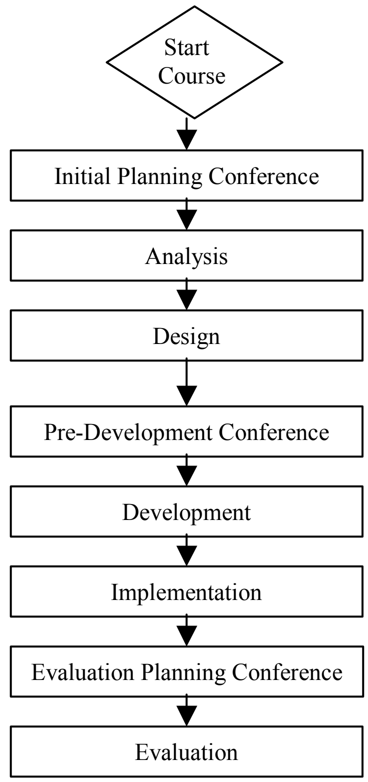


Figure 4. DLTD Development Process

a. Start Course

As would be expected, the initial step is to start the course. This decision is not made in a vacuum. As stated earlier, there are a limited amount of course development teams, and a multitude of course materials for them to work on. Each team is responsible for a certain mix of MOS's and the accompanying existing courses, as well as the need for new courses in those specialties. Additionally, the team is responsible for certain modules of the PME courses, and a number of PSTs. A careful look is taken by the supervisory element of DLTD, MCI headquarters, and well as within the team itself to determine what product will actually be worked on. Additionally, input is taken into account from the Marines of the Fleet Marine Force, Marine Corps Training and Education Command, and from the heads of the MOS's themselves. All of these factors together determine the optimal course of action. Once the decision is made that work exists and needs to be done, the key players get together to map out a strategy.

b. Initial Planning Conference

The Initial Planning Conference at MCI brings all of the key players together to start off on the right foot. Representatives from the Course Development Team, editors, layout specialists, and logisticians are present to come up with a good initial plan. This plan includes a start date, milestones for each following step in the process, and at the end a project plan that will be used to track the progress of the project. Once this is agreed upon by all hands, progress continues to the next step.

c. Analysis

The analysis phase is characterized by project team doing a detailed study of the subject at hand, and determining what key topics will be presented to the students. Additionally, these topics are discussed with Subject Matter Experts (SMEs) as well as with the formal educational bodies such as the formal schools for the MOS, and the Training and Education Command at Quantico. Once everyone with a stake approves the content, the next phase is undertaken.

d. Design

The design phase is a relatively short one. The course development team takes the input from the analysis phase, and places it into a specific plan of what the finished product is going to look like. The course is broken down by study unit, choices of media are made, and the general layout is established. With this specific roadmap, the team is now ready to progress to the next step.

e. Pre-Development Conference

Once design is complete, the players from the IPC are again brought together to verify the work that has taken place since their initial meeting. Milestones are fine-tuned to ensure that the production will proceed on track. At this point, the actual development of the course commences.

f. Development

The meat and potatoes of course development is conducted during this phase. The Distance Learning Instructor puts pencil to paper and writes the actual course content. Additionally, the course developer works hand in hand with the graphics and layout specialists to optimize the use of graphs, pictures, and tables to best illustrate the subject matter. As well, the Instructional Systems Specialist (ISS), a civilian expert in education, assists the course developer to ensure that the material is written in a way to actually teach the student.

Courses are usually broken down into logical segments called study units. The developer, upon completion of a study unit, allows the ISS and the Project Manager to review and amend his work. An internal meeting between the team resolves any differences, and the work is then submitted to an editor for proofreading. Upon completion of all of the study units for a particular course, and appropriate examination materials, the course is the validated. This entails taking the course to its intended student population, and having them take the course in a short period of time, and gathering feedback to make last minute corrections. Once all late revisions are made, the course is made ready for enrollment.

g. Implementation

During the Implementation phase, the course is sent out of the department to be printed. Once the materials are on hand and ready for issue, the course is placed in MCIAIS and enrollments are taken. Students take the course, and to receive credit, send in their completed examinations for grading. Once a passing grade is achieved, credit is given in MCIAIS, and the student is credited with completing the course.

h. Evaluation Planning Conference

The Evaluation Planning Conference is for all intents and purposes an after-action meeting of all of the key stakeholders. They evaluate how the development process worked in the case of the newly developed product, and look for obvious shortcomings. They look not only at the quality of the product, and the procedures used to get the product completed, but also at any development delay, and the cause of that delay. The end state of this meeting is a list of what went well and could be implemented in other projects, and what failed, and how to avoid those pitfalls in the future.

i. Evaluation

Evaluation is a complete look at the educational and technical validity of the current material, and how it was produced. If the changes required are of an extensive enough nature, then the course is considered as a revision, and begins the process for course development at the IPC stage. If only minor alterations are required, then a small chance without the process is done, and the materials are reprinted. If the material is so outdated and no duplicate materials are required, such as the phasing out of a piece of equipment, then the course is deactivated, and no replacement is produced.

4. Course Maintenance

A less formal process than development exists within DLTD known as course maintenance. Existing courses are broken down into groups, roughly correlating to MOS, which are administered by Distance Learning Instructors and Project Managers. Each of these courses must be maintained, as there are students who are actively enrolled in these courses. The DLI is the key link in this chain. He must monitor stock on hand, and request reorder of materials when his stock is running low. Additionally, he receives all student feedback concerning his courses. He notes any possible discrepancies, and looks for items in the feedback that can be of value in any upcoming evaluation of that course.

The DTI interfaces directly with MCIAIS through the R-94 process. This process, named after the old paper form R-94, is making any updates to courses administered in the database. Information such as which version of the course is active, study hours required for the course, and Reserve Retirement Credits allowed, are contained in MCIAIS in the Course Program Description table. Fields that can periodically change and therefore be updated, are modified by the DTI, and tracked within MCIAIS.

The function of maintenance, though becoming more integrated into the world of information technology, is often an afterthought, and is an area that is ripe for an IT optimization solution.

D. INFORMATION TECHNOLOGY DEPARTMENT (ITD)

As its name would indicate, the Information Technology Department is responsible to the Institute for providing IT support for the development, issuing, grading, maintenance, and record keeping of distance learning materials. Additionally, it is charged with developing better practices through the use of IT, maintaining the web capability of MCI, maintaining interface with the Marine Corps Total Force System (MCTFS), and keeping the MCI and Marine Barracks Washington network up and running. As with all IT departments, it has a very large task set, limited resources, and an expectation of 100% uptime. Nothing is new here. ITD is tasked with bring the data warehouse up, and educating the users to its potential, and teaching them to get the most out of it. Once the initial set-up tasks concerning the data warehouses is complete, then ITD must be vigilant to ensure that it is used to the utmost of its capability.

1. Networking

As stated above, one of ITD's major functions is to maintain the network. This is deceiving, as there are multiple pieces of different networks that must be maintained.

a. Marine Barracks Washington Network

As stated in Chapter II, MCI is an integral part of Marine Barracks Washington. As this post is geographically separate from any other Marine Corps facility, it is inherent that the command maintain its own network. Since MCI has a complete department devoted to IT, and the command Marine Barracks Washington has no such capability, it falls to MCI's ITD to maintain the entire network. What makes this additionally challenging is that there is a small physical separation between MCI and the Barracks headquarters, and they are separated by a major highway. Additionally, sub-elements of the command exist as far away as Camp David, MD, and the United Stated Naval Academy in Annapolis, MD, at distances of over fifty miles. As with most non-IT network users, little understanding of network connectivity is held by the Barracks, but maximum performance is expected.

A major part of this network as a whole is MCI itself. Fortunately, no geographical separation exists, but many computers of differing needs are present. Course materials are developed, maintained, and administered on this network, and it seems that every worker within MCI needs his or her own workstation.

b. MCI Online

MCI's students' means of tracking their own course progress and gaining information from MCI exists on MCI online, MCI's website. This portal, which is entered through the user entering his last name, social security number, and birthday, has allowed the student access to his own educational records. It is also interactive in that it allows Marines to enroll in courses. Keeping this site up, as well as ensuring that its interface with MCIAIS, its back-end database, is correct is a challenging task for ITD. Of additional difficulty is that students are deployed world-wide, and hits to the site are common in the middle of the night. That is because the middle of the night in Washington is the middle of the day in places like Okinawa and Iraq. This makes windows for maintenance difficult to find, and a constant load being applied to MCIAIS.

c. MCIAIS and MCTFS

MCIAIS will be looked at in detail in the next chapter, but suffice it to say that maintaining a 50+ gigabyte database with constant hits from the website is daunting in itself. MCIAIS also must remain connected to the Marine Corps Total Force System. This larger Marine Corps-wide database, maintained in Kansas City, is the main administrative database for all Marines. Pay, entitlements, and promotion information are kept in this data reservoir. As stated in a previous chapter, completion of MCI courses is critical to Marine's promotions. If the completion is not recorded in MCTFS, then for the purpose of promotion, the Marine has not completed the course, even if it is in MCIAIS. Maintaining this link to ensure transferal of this critical information is a key responsibility of ITD.

d. Computer Support

As is expected in IT in any organization, support of the individual user, whether it is in setting up workstations, establishing accounts, setting passwords, troubleshooting hardware, or monitoring Internet usage are functions that must be accomplished. The networking section of ITD performs all of these functions, and as with most IT departments, spends most of its time in this area.

e. Extracting Knowledge from the Data Sources

A crucial component of the IT department's job description is the development of knowledge from MCI's data sources. Currently, a four-step process is used to get the users information that they need to make intelligent decisions concerning business at the Institute.

First, the requesting agency, a worker in another department, approached the IT professional with what question they would like to have answered by the data source. This could range from how many students are enrolled in a particular course to how many 21 year-old Lance Corporals stationed at Camp Lejeune failed version B of course 08.20 during the month of August. Once the question is ascertained, then the IT professional develops the query for the data source, runs the query and presents the results to the requestor in a format that they can understand and use.

This particular task is one of the key reasons for the implementation of the data warehouse. In the current MCIAIS environment, several obstacles stand in the way to an efficient completion of this task. First, MCIAIS is always online to support MCI Online and MCTFS. Querying the online database is a process that causes the entire system to run significantly more slowly. Second, as will be described in Chapter V, MCIAIS has a very complicated data structure. To write the query in standard Structured Query Language (SQL) is often very challenging, if possible at all. Once the query is run, then formatting is completed and the user has his answers. This specific set of tasks will be analyzed in its current configuration using MCIAIS and then compared to the same process using the data warehouse to determine the return on knowledge differential in Chapter VII.

2. Programming

The programming section within ITD's mission is less all encompassing than the networking section's mission. Their function is to fix identified software problems, and to provide new solutions for the command through developing code. MCI is constantly striving to improve its support to the ultimate customer: the MCI student. By finding innovative solutions to the problem of providing distance education, the programming section within ITD fulfills its charter. This section completed the initial development and launch of MCI Online, and a large portion of the implementation of the Data Warehouse was also within their purview.

E. STUDENT SERVICES AND SUPPORT DEPARTMENT (SSD)

This department had a many functions, and by far the most Marines at its disposal. Its two main functions are customer support and logistics. In both areas, SSD can improve its operating efficiency through use of the warehouse.

1. Customer Support

The customer support division has two main functions. One is in course administration, and the other is in customer service.

a. Course Administration

Within administration, students are enrolled in courses, and their exams are graded. Enrollment is a function that has largely become automated. Over 90% of students now enroll in courses via MCI Online. Since some students do not have the ability to enroll online, the customer services division maintains the capability of accepting enrollments by phone or by mail. The division then performs data entry into MCIAIS for the student. Exam grading is still done with paper examination forms being manually inserted into grading machines. The examination results are then entered into MCIAIS. If a student fails to achieve a passing score, the Marines of customer support enter this data, and ensure that an alternate examination form is sent to the student for another chance at passing the course. Disenrollment, or being dropped from a course used to be a function of this division as well. With the integration of MCIAIS, and fixed time limits on course completion, this process is completely automated at this time.

b. Customer Service

Customer service consists of a call center, which includes both Marines physically answering telephones, and more commonly answering e-mails. These Marines answer all manner of questions concerning MCI. The answer the mundane information readily available on the website such as courses available, how to enroll, and specific information about courses. More commonly, they answer specific questions concerning a student's progress in a course. These questions are most often how much time is available to complete this course, have you received my examination yet, have I passed, is the grade in the system, and will the completion count towards my next promotion.

An additional customer service entity in place is the PME Help Desk. Because PME courses are often much longer and more complex, and affect the promotion of senior enlisted Marines and Officers, a special section was created to deal with this group of courses. The same types of questions are asked, as well as a litany of other more complicated PME completion inquiries. Normally, more experienced Marines, not easily frustrated by answering questions of senior Marines are paced at this workstation.

2. Logistics

This multi-function division used to be a department unto itself due to the multitude of tasks it performs, and the sheer magnitude of its operation. The main functions of this division are printing, contracting, warehousing, and postal operations.

a. Printing

Although the capability within MCI's facility are limited, printing operations of examinations and of smaller courses are conducted in house. Additionally, when time is of the essence, like producing a course on Middle Eastern Military Forces before and impending war, or a new version of a PME exam is to be distributed prior to the convening of a promotion board, the printing shop can handle these orders.

b. Contracting

Contracting at the higher level is done at MCI's headquarters. The contracting done by the logistics division is establishing contracts to print large quantities of course materials. Due to the above stated limitations of the in-house printing shop, many jobs are sent to commercial printers. The contracts for these as well as accountability and quality control for the materials are the responsibility of logistics.

c. Warehousing

Maintaining an adequate supply of materials for issue of courses and examinations is a very large task. Keeping the count accurate both in MCIAIS as well as in the warehouse is critical to the process as a whole. Failing to monitor stock, or to allow the counts between what is actually on hand and what is in the database can be very costly. Special print orders for courses that are not on hand can be very costly, and upset the very tightly controlled MCI budget a great deal. MCI recently implemented a bar coding system to manage its warehouse, that is fully integrated with MCIAIS. This has increased efficiency dramatically.

d. Postal Operations

Until online courses and examinations are available for all courses, a robust postal system must be maintained at MCI. Be the course books, examinations, or IMIs, the materials need to be sent to the students upon enrollment, and the examinations need to be delivered to the graders. As recently as five years ago, it often took two months for a course to actually get in the hands of the students. An automated postal system within the warehouse, coupled with faster enrollment procedures has decreased that time to a fraction of what it once was.

THIS PAGE INTENTIONALLY LEFT BLANK

V. MARINE CORPS AUTOMATED INFORMATION SYSTEM

A. THE DATABASE

The Marine Corps Automated Information System (MCIAIS), previously referenced frequently in this document, is the one all-consuming database used by MCI to conduct all of its operations. This database is the target of the warehouse, and its migration to the warehouse platform is critical to the success of the project. MCIAIS is a 50+ gigabyte Oracle 8i database that is running on a Compaq Proliant ML530 Pentium III Xeon system. It has over 90 different tables, each containing a multitude of fields. As can be imagined, given the size of the database and the complexity of MCI's operation, the entity-relationship diagram of this database looks like the roadmap of a major metropolitan city. This is depicted below in Figure 5.

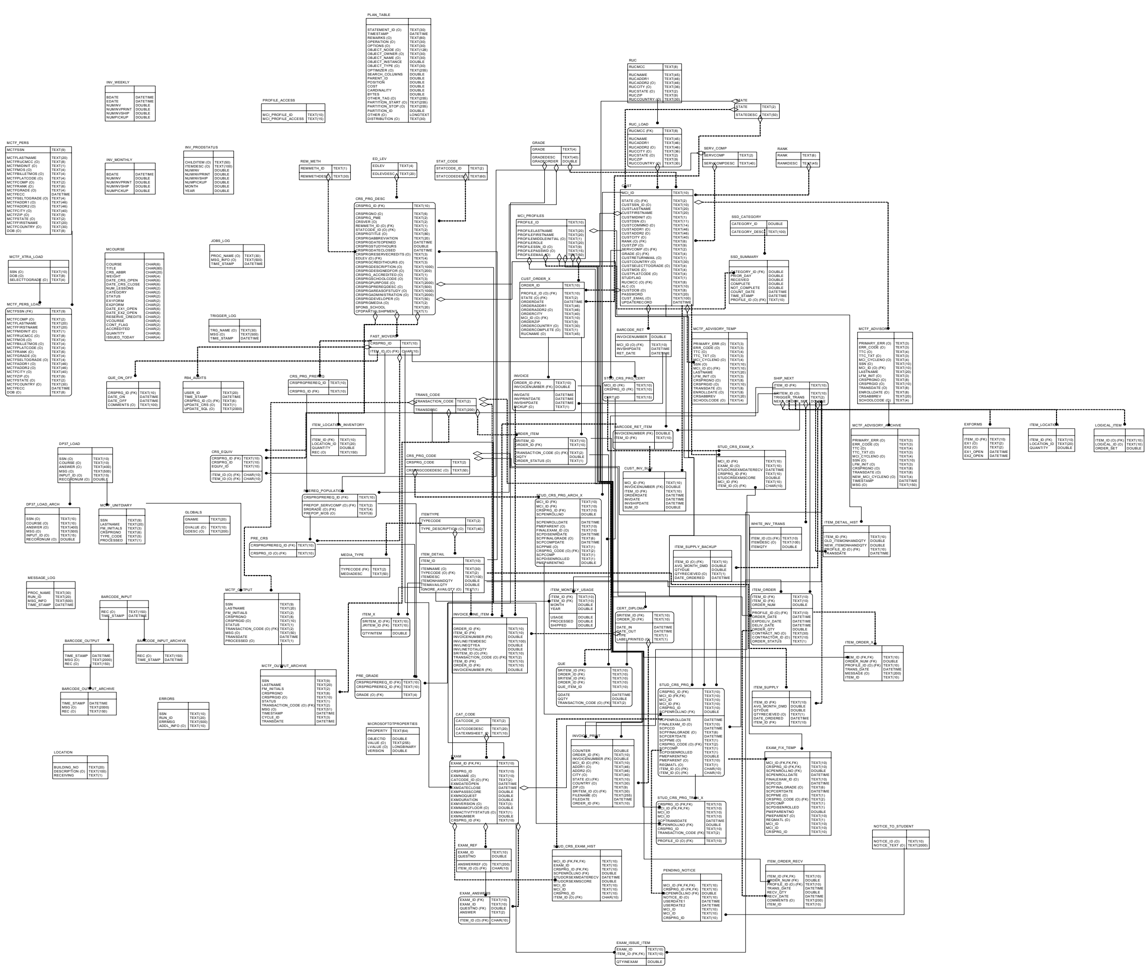


Figure 5. MCIAIS

Like many databases of this kind, MCAIS has evolved over time. Initially, student records, course information and logistics data were stored as separate data entities. With new technology, and more powerful database software at MCI's disposal, the data of all of the "homegrown" databases were merged into one "super database": MCIAIS.

B. SIGNIFICANT PARTS

It is obvious that to get a good feel for this database, it must be broken down onto smaller chunks to get a good understanding of it. It must be understood by looking at the complexity of Figure 4 that not all fields and relationships can accurately be portrayed using a chunking methodology. What follows is a look at some of the key sections, and an explanation of the important elements and their relationship to the whole. The starting point is as it should be, with the customer.

1. Student Information

As it is to be expected, MCI exists to provide a service to a customer: a student enrolling in and taking courses. To be a successful enterprise, MCI must ensure that it is servicing its student's needs. Figure 6, shown below is a slice of MCIAIS that primarily focuses on the customer.

As is to be expected, the critical table in this area is Customer (CUST). No less than eleven other tables either feed or are fed by this table. Its primary key, MCI_ID, is the locally generated number that keeps all of the students in sequence. Critical elements of this table include all of the critical information about the student: name, rank, address, MOS, and information that allow the student to interface with MCI online such as password. Tables feeding into CUST include STATE, RANK, SERVCOMP, GRADE, and RUCMCC. The purpose of these tables is to allow for the dropdown menus when filling in student information.

Critical elements connected on the outbound side include links to the Marine Corps Total Force System (MCTFS), student course progress, customer orders, invoices and barcodes, and student profile tables.

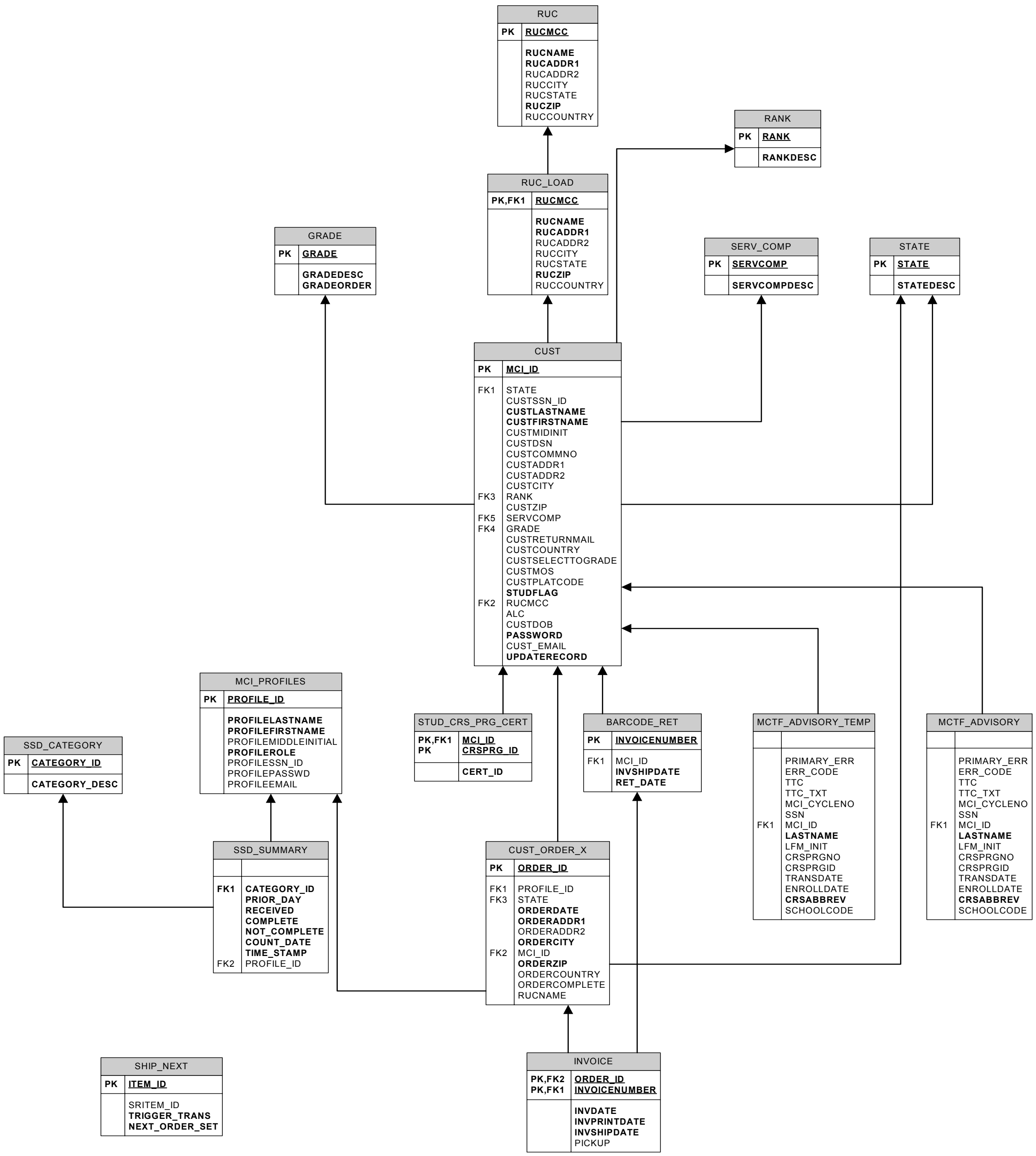


Figure 6. Student Information Section

2. Course Information

The next piece of the puzzle in MCIAIS is information concerning the courses themselves. A slice of these tables is shown below in Figure 7.

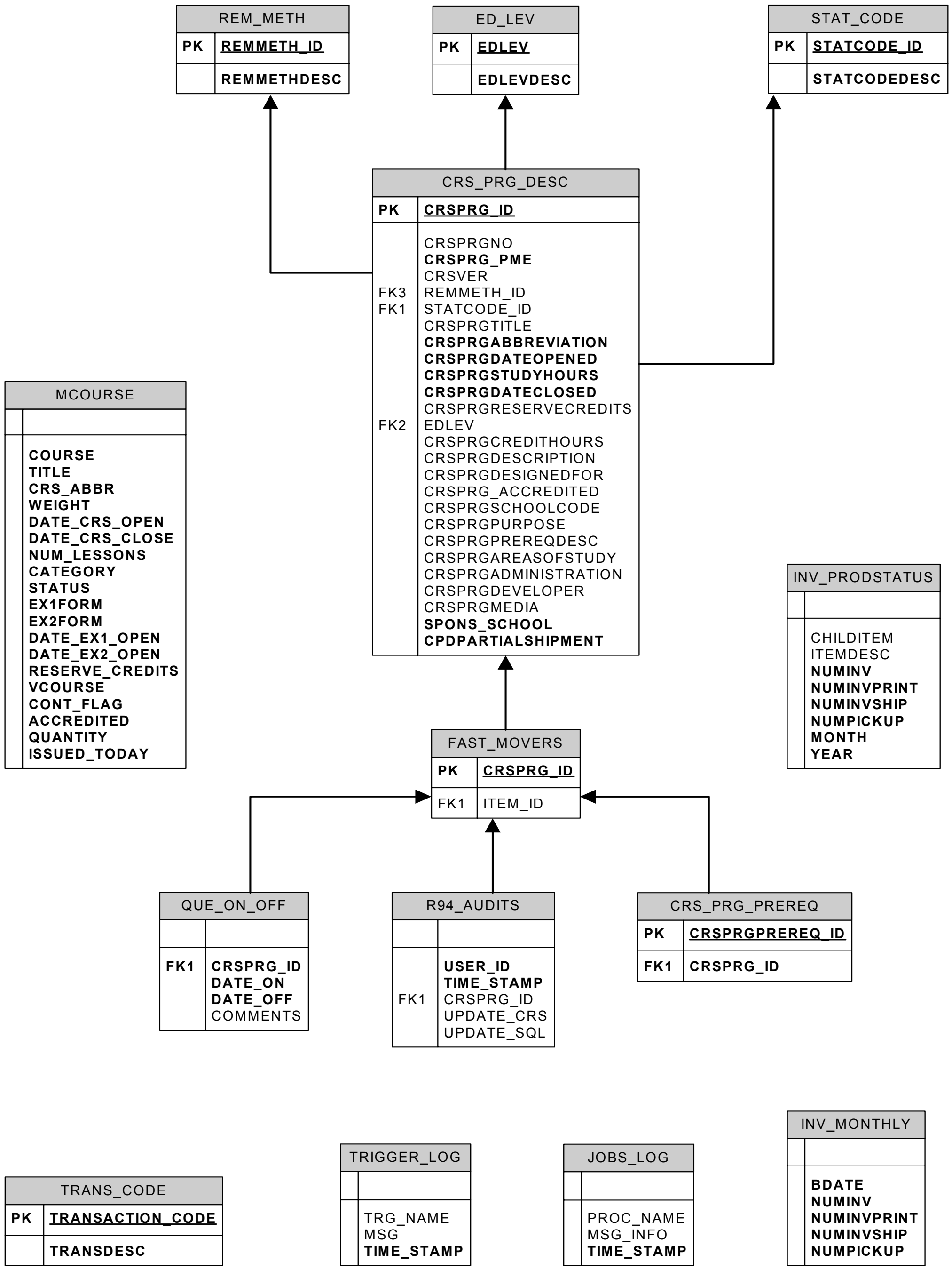


Figure 7. Course Information Section

The information contained in the table Course Program Description (CRS_PRG_DESC) is the meat and potatoes of the course section. It contains critical elements about each course such as title, program number, date opened, study hours credit obtained, education level, etc. It is fed by several tables such as Education Level and Status Codes, and feeds a critical table, Fast Movers. This table is used by MCI to keep track of those courses that see much greater enrollment, so to keep adequate stock on hand and to make life easier in the warehouse. The table R94_AUDITS is also a key table in this section, as it the table that the DLTD developers and managers do direct entry into to make changes to the course profiles, as is consistent with their job descriptions.

3. Course Ordering Information

Now that we know how we are keeping track of the students and the courses, it would be wise to look at how the student and the courses get together. The critical table in this section is the table CUST_ORDER_X. This table is fed either by an MCI staff member entering the information on site, or by the student himself using the web interface, MCI Online. The table basically is who the student is, where he is located, and when he placed the order. The who part is drawn from MCI_PROFILES, part of the customer information piece. An invoice number is then generated for tracking and for use by logistics. This also links into a field called BARCODE_RET, which is used by logistics in their tracking and shipping plan using bar-coding. A link exists as well to the course information side, so the system is keeping track of what course is ordered and by whom. A more detailed drawing is represented below in Figure 8.

This section is critical in that it links together three important elements within the bigger picture: the student who is enrolling in the course, the course that is being enrolled in, and that this enrollment is being chronicled using an invoice as a device. In this sense, MCIAIS is very similar to any e-commerce database on which online transactions are taking place.

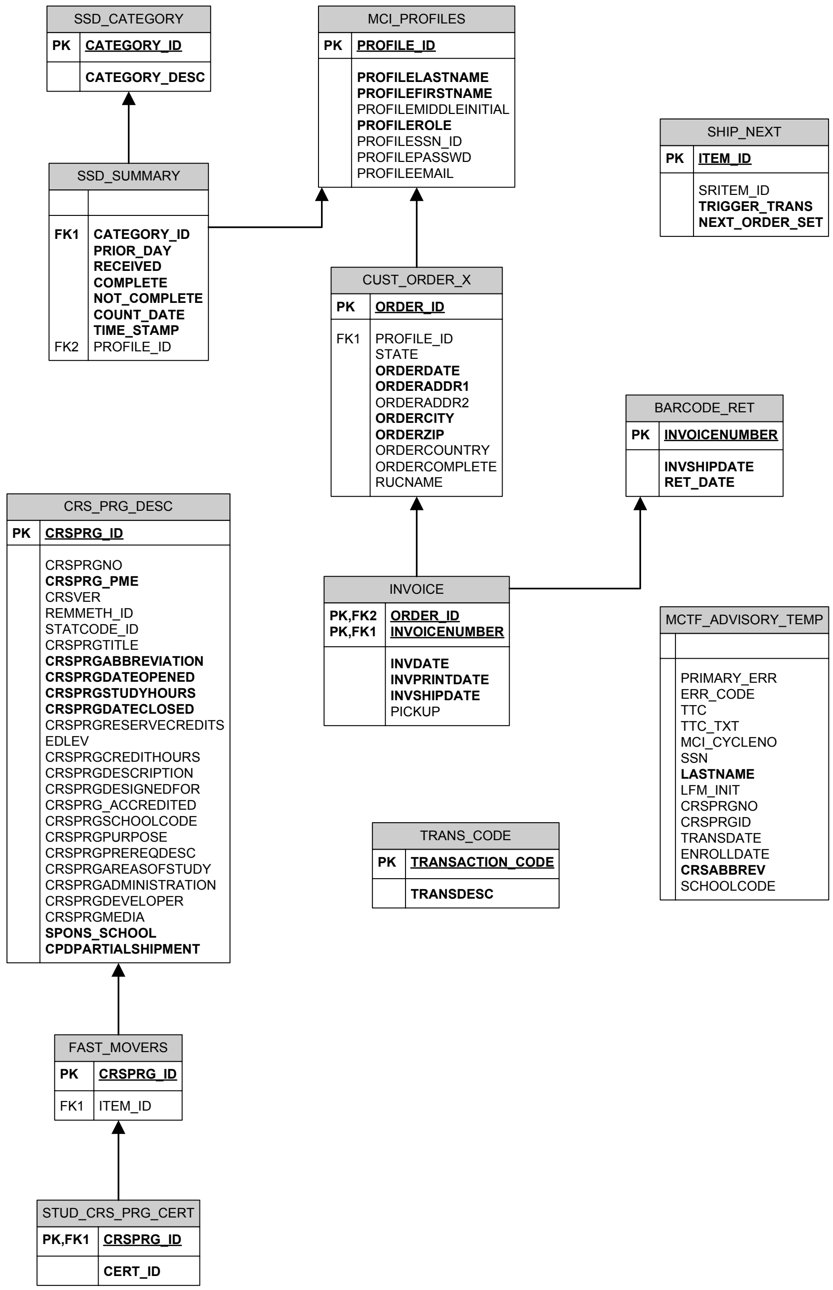


Figure 8. Course Ordering Section

4. Logistics Information

After the student has done his part and enrolled in the course, it is than up to MCI to ensure that he gets the requested materials in order to take the course. This is up to the logistics section within SSD. Details are below in Figure 9.

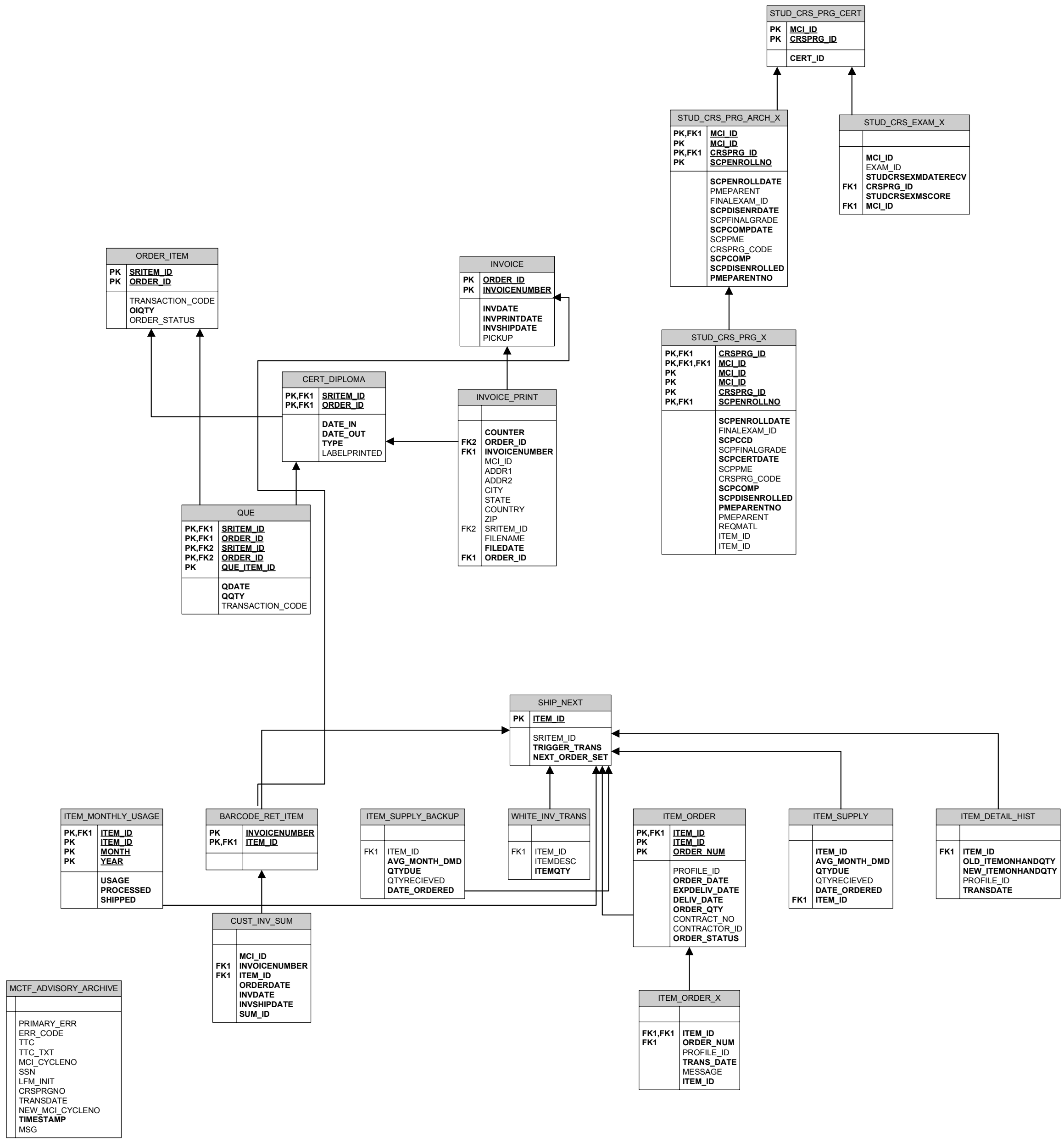


Figure 9. Logistics Section

This section outlines the relationships between the order done by the student, its scanning into the system using the barcode, and the logical link to shipping data. Once a course is shipped, MCIAIS will automatically decrement the supply on-hand with links to the item supply fields resident in the database. If all works well, MCI puts the requested materials in the hands of the student, tracks that transaction, and keeps inventory control as well.

5. Examination Information

Once MCI gets the instructional materials to the learner, it is the task of that learner to complete the course of instruction. This is done through the administration of an examination. The student completes a multiple choice, scannable examination form at his location and sends that form back to MCI for grading. Relevant portions of MCIAIS are shown below in Figure 10.

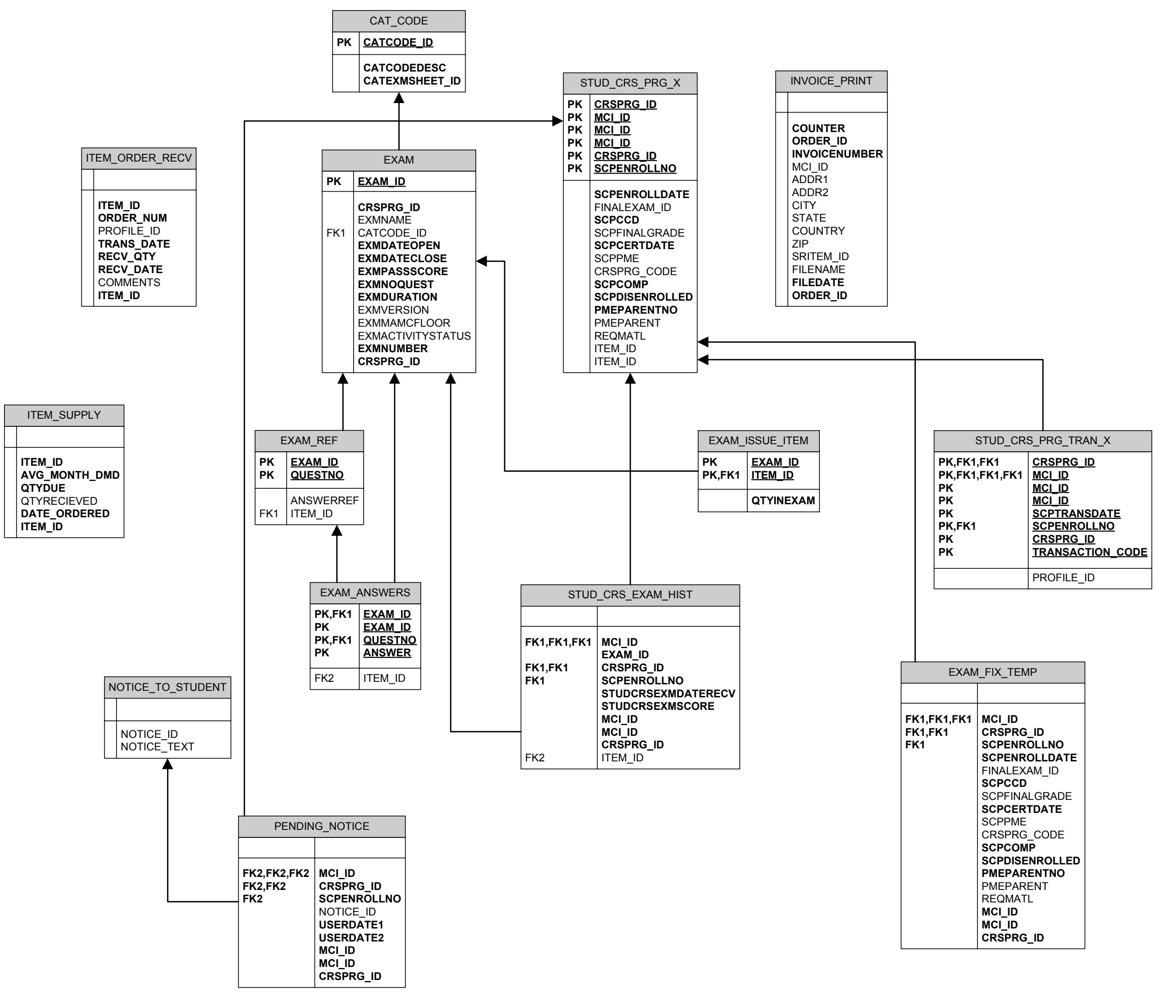


Figure 10. Examination Information Section

Once the examination has been completed by the student and received by MCI, it is graded by SSD using the EXAM, EXAM_REF, and EXAM_ANSWER fields using the scanning tool as an interface. This information is entered into the student's course exam history and one of two things happen. If the student has achieved a satisfactory score and passed the course, the completion is logged and a course completion certificate is generated, as well as the student's history is updated to show completion. This then links to MCTFS for promotion purposes. If on the other hand the student fails, a notice is generated to the student of failure, an alternate exam is sent, and the exam history is updated.

C. OTHER ASPECTS OF MCIAIS

The preceding section is the 90% solution to the database, but it doesn't tell the whole story. In addition to tracking individual students, groups by location, unit, rank, and many other classifications can be done. Additionally, significant functionality had to be built in to MCIAIS in order for it to link properly with MCTFS. Without this linkage, students' records would be complete at MCI, but they wouldn't get the credit where it counted: in the Marine Corps administrative system.

As with many databases in use by entities such as MCI, tables and relationships had to be built to satisfy the needs of the executives. These tables often don't make sense from the database administrator's point of view of having a clean entity-relationship diagram of his database, but they are what make requested queries and reports more expedient. This by no means makes MCIAIS user-friendly for this purpose. MCIAIS is built for the functionality of tracking data, enabling MCI Online and linking with MCTFS. Its use as a tool to assist in decision support and data analysis is in fact a very low priority. To truly capture that functionality, the data must be reordered and placed in an entirely new format to unlock the potential wealth of such an implementation. That is the reason for the investment in the data warehouse.

At this point, it is time to start looking at the actual implementation of the data warehouse at MCI.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. IMPLEMENTATION OF THE WAREHOUSE

The implementation will be looked at in two phases. The first phase is in setting up of the hardware and software, extracting, transforming and loading of the data from MCIAIS, and the replication of existing reports and queries in the warehouse environment. The second phase will look at using the capabilities of the warehouse to actually achieve some value-added.

A. SETTING UP THE WAREHOUSE (PHASE I)

Initially, MCI had a vision of data being analyzed and mined form their as yet untapped source of MCIAIS. Logical thought told them that data warehousing might offer the solution that they were looking for. Choice selection of brands was not difficult as MCIAIS is an Oracle 8i database, and MCI already had the appropriate licenses and personnel familiar with Oracle products, so an Oracle solution was obvious. The choice made was to go with the Oracle 9i warehouse configuration. This selection minimized costs, and made implementation much easier in the long run.

1. Hardware

The hardware by itself was not that extensive. Two Dell P6600 servers, one with eight disks in the array, and the other with four disks in the array were purchased and installed. Although not the originally ordered configuration, MCI successfully installed the software on the servers and brought the systems up.

These two servers where then connected to the MBW local area network via the existing Fibre Channel Gigabit switch using Cat 5 cable at a data rate of 100 megabits per second. This switch is than patched into the LAN giving connectivity to the user. A schematic is shown below as Figure 11.

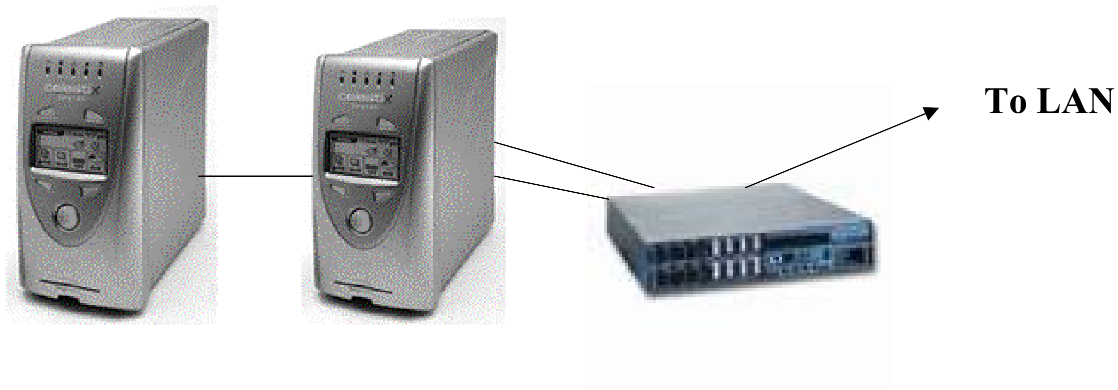


Figure 11. Physical Warehouse Setup

2. Software

Software included the Oracle 9i warehouse suite, which includes the Warehouse Builder. The Warehouse Builder is the core of the system that enables the other pieces to accept information. In addition to the Warehouse Builder, Oracle 91 AS Discoverer is used performs extensive metadata reporting and integration, and Oracle Workflow and Oracle Enterprise Manager are used to enable the system. All of these pieces were loaded to the new servers in order to bring the warehouse to operational capability, and get it running.

3. Extracting, Transform, and Load (ETL)

ETL was performed using the automated ETL functionality of Oracle 9i Warehouse Builder. Being that MCIAIS is the lone source data repository, and that it is an Oracle database, Warehouse Builder was not taxed to the limits of its capabilities in the ETL process. Warehouse Builder is designed to integrate many disparate data sources such as flat files, mainframes, spreadsheets, other types of conventional databases, as well as Oracle databases. Moving one Oracle 8i database via this tool was rather easy.

During the ETL process, Oracle Warehouse Builder "grabbed" the data from MCIAIS from multiple columns, and in some cases entire tables were extracted. This initial extract was based on the current reports and queries that MCIAIS provides, which are detailed below.

Once extracted, the transformation process begins. At this point the extracted data is reformatted, or transformed into new tables that relate in slightly different ways than the production database. This transformation allows for much easier access to the data, in the way that the entity will use it. This does away with the complicated SQL queries that are required to produce the same reports in MCIAIS. The last chapter clearly illustrated the complexity of MCIAIS; in the warehouse environment, queries and reports of interest are much more easily designed and produced due to the data transformation.

Once the extraction and transformation are complete, the reformatted data is loaded into the target: the data warehouse. Another function of this process is to "cleanse" the data. As shown in Chapter V, MCIAIS has a multitude of tables with even more relationships. Although MCI has remained vigilant to keep MCIAIS as clean as possible, some duplications of data, as well as orphaned tables exist to make the interface with the web or MCTFS run properly. These aspects of MCIAIS are left behind, taking only those data elements needed for the warehouse functionality. What is loaded on the warehouse servers is a much "cleaner" version of MCIAIS, with only the information needed to improve core processes.

4. Replication of Existing Reports

Before the construction of the warehouse, MCI used normal structured query language (SQL) queries of MCIAIS to develop a family of management reports. These reports were available on the MBW Intranet for use by decision makers across the breadth of MCI. What follows below is a summary of the reports as they existed from the live feed.

a. Daily Enrollments Report

This report is an updated view of what course enrollments have taken place on a particular day. It lists by course, how many new enrollments are taken, and breaks them down by number enrolled via the website, and how many were manually entered at MCI (taken via phone or were mailed in by the student). This is primarily used by logistics to keep track of stock, as developers and executives need to see a greater time period to make good analysis.

b. Monthly Enrollments Report

This report offers the same information as the daily report, but also rolls up how many total courses were enrolled in over the month's time. This report gives the decision maker a better grasp of how popular courses are, what kind of restocking is necessary, and helps to give indications that certain courses need to be weeded out of the inventory. Currently, the total monthly enrollments MCI-wide average about 50,000.

c. Enrollments (All)

Functionality using this option includes letting the user search specific courses, geographic areas, rank, or other criteria to get a better grip on how enrollments are going over a specified period of time. This report looks at all enrollments, both MOS related and PME.

d. Enrollments (PME-adjusted)

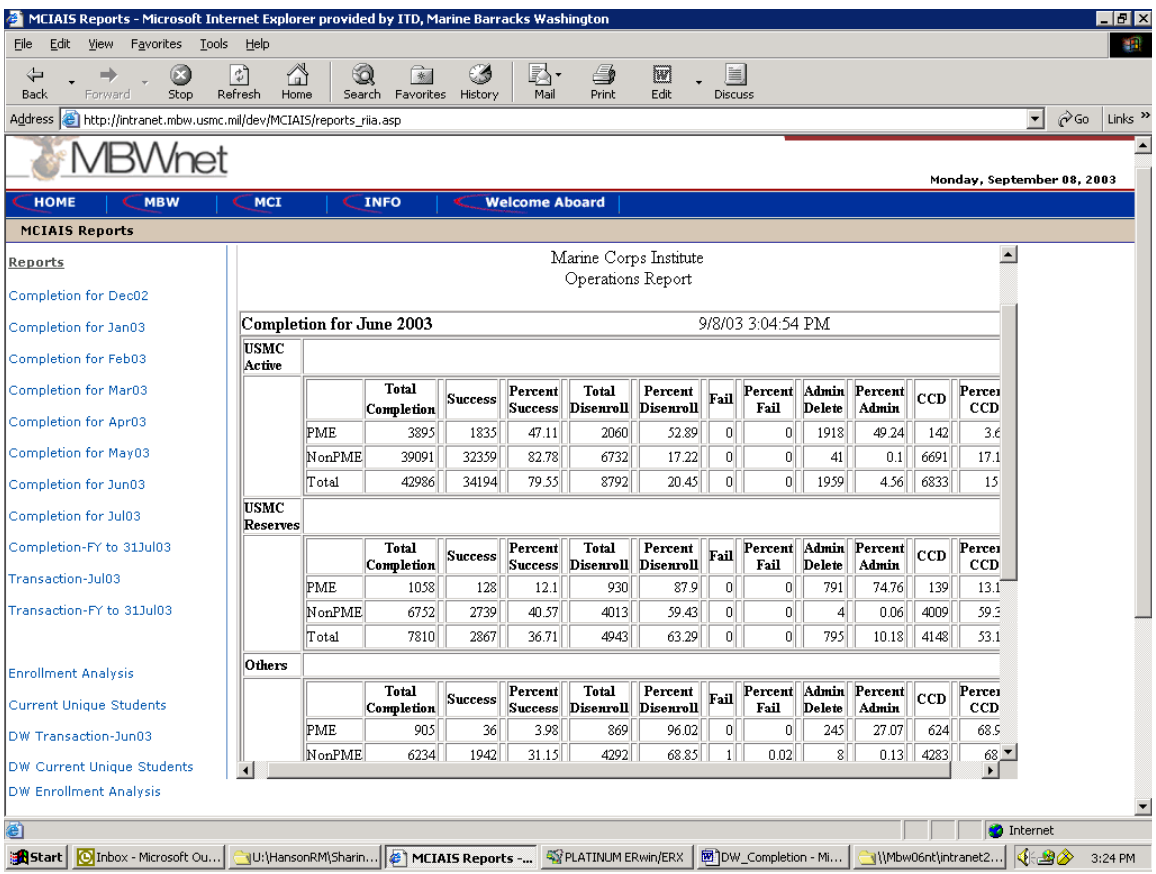
This report has identical functionality as the previous report, only it filters out PME enrollments. Due to the large number of PME enrollments due to their direct link to promotion, it is sometimes more accurate for DLTD to look at all of the course materials except for PME, because this is what they primarily develop and manage.

e. Monthly Completions

This report rolls up on a monthly basis how many individual completions have occurred. This gives MCI a good read of how the student's progress is coming, how many diplomas are needed, and trend data of what time of the year a larger portion of completions are coming in.

f. Completion Percentage

This report gives a percentage by course on completions. This allows management and developers to look at how many of the students actually are finishing their course of instruction. If the percentage is too low, it could be due to course material that is too difficult, exams that are too hard, or that the material just doesn't hold the student's interest. If completion is too high, then the course and exams may be too easy. This table as it is pulled from MCIAIS is shown below in Table 3.



		Total Completion	Success	Percent Success	Total Disenroll	Percent Disenroll	Fail	Percent Fail	Admin Delete	Percent Admin	CCD	Percer CCD
USMC Active												
	PME	3895	1835	47.11	2060	52.89	0	0	1918	49.24	142	3.6
	NonPME	39091	32359	82.78	6732	17.22	0	0	41	0.1	6691	17.1
	Total	42986	34194	79.55	8792	20.45	0	0	1959	4.56	6833	15
USMC Reserves												
	PME	1058	128	12.1	930	87.9	0	0	791	74.76	139	13.1
	NonPME	6752	2739	40.57	4013	59.43	0	0	4	0.06	4009	59.3
	Total	7810	2867	36.71	4943	63.29	0	0	795	10.18	4148	53.1
Others												
	PME	905	36	3.98	869	96.02	0	0	245	27.07	624	68.9
	NonPME	6234	1942	31.15	4292	68.85	1	0.02	8	0.13	4283	68

Table 3 Completion Percentage

g. Current Unique Students

This is a report giving a readout of how many individuals are enrolled in at least one course, filtering out multiple enrollees. This gives good demographic information of who the customers actually are.

h. Find Student

This search function allows the user to locate a specific student and see the report of what that student is enrolled in and what progress is being made. This assists in helping students who cannot get a good answer from the website and need further assistance.

i. Find Reporting Unit Code (RUC)

The report allows the user to search MCI activity by Reporting Unit Code (RUC), or numerically by unit. If a bulk enrollment is done by a unit this report works well to track information. Also, it allows the user to search a unit for information if it cannot be otherwise accessed by student identifier.

j. Course Data File

The Course Data File is information provided by course rather than by student or unit. This functionality helps those who manage specific courses to gather information about their student population.

k. Course Information Report

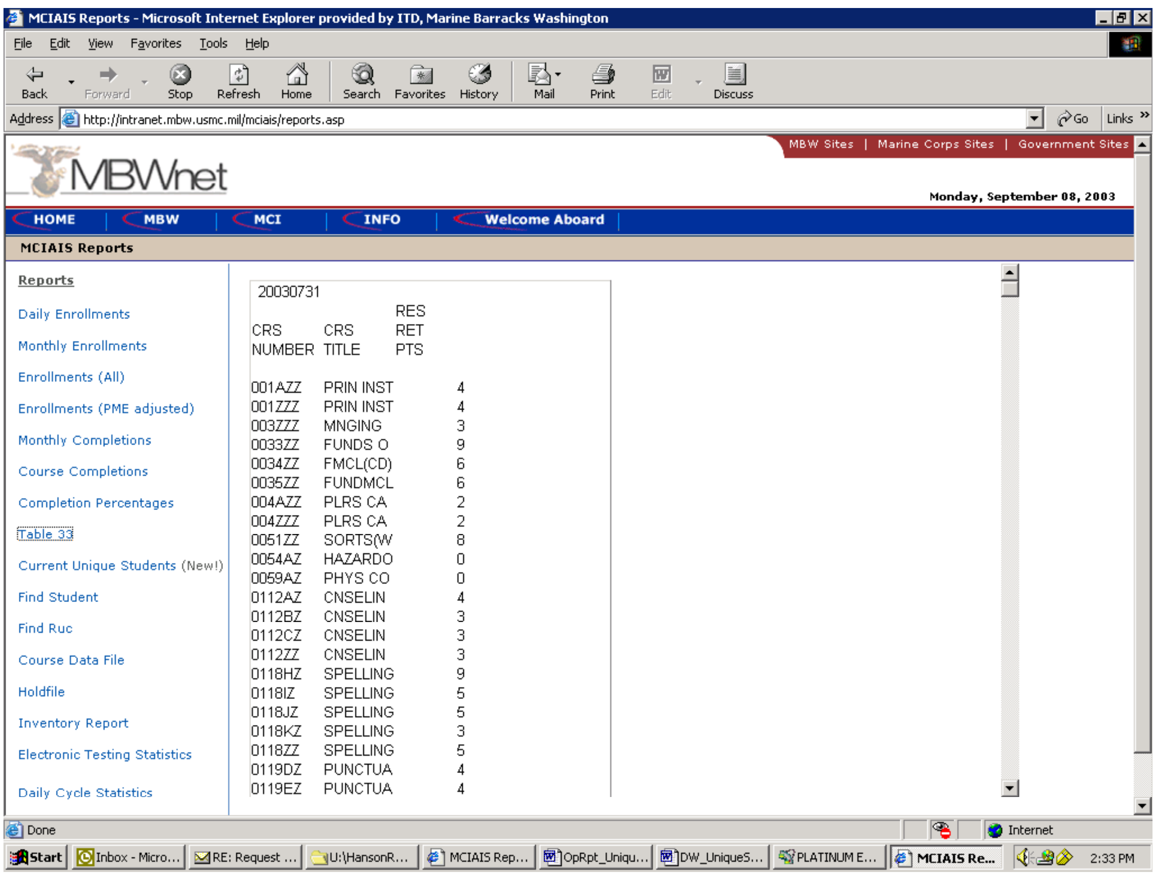
This report is not so much about enrollments in a particular course, but is information about the specific course. Items in this report include title, basic description, study hours, reserve retirement credits, etc. This is good information to have at one's fingertips when evaluating and managing what work needs to be done in the time ahead.

l. Marine Test Report

This report deals with course examinations specifically. Rather than getting generalities from completion reports, this report provides specific information of exams taken by specific students.

m. "Table 33"

Table 33 is a custom generated report containing all course numbers that have existed and their version letters. The corresponding course titles, and the reserve retirement credits associated with each version. This is used as a good historical archive of what has taken place at MCI in the past, as well as what is current. This table is shown below as Table 4.



20030731

CRS NUMBER	CRS TITLE	RES RET PTS
001AZZ	PRIN INST	4
001ZZZ	PRIN INST	4
003ZZZ	MNGING	3
0033ZZ	FUNDS O	9
0034ZZ	FMCL(CD)	6
0035ZZ	FUNDMCL	6
004AZZ	PLRS CA	2
004ZZZ	PLRS CA	2
0051ZZ	SORTS(W	8
0054AZ	HAZARDO	0
0059AZ	PHYS CO	0
0112AZ	CNSELIN	4
0112BZ	CNSELIN	3
0112CZ	CNSELIN	3
0112ZZ	CNSELIN	3
0118HZ	SPELLING	9
0118IZ	SPELLING	5
0118JZ	SPELLING	5
0118KZ	SPELLING	3
0118ZZ	SPELLING	5
0119DZ	PUNCTUA	4
0119EZ	PUNCTUA	4

Table 4 Table 33

B. RESULTS OF PHASE I

Phase I completion gives MCI a functioning data warehouse. Its data source for analysis is the reconstructed "clean" database, which is not tied to the online web application. This gives MCI the opportunity to conduct analysis without having to slow down MCIAIS and its online functionality, or sever the connection with MCTFS. Because of this, MCI has achieved a 333% increase in the speed in which reports and queries can be processed as opposed to pulling them from MCIAIS. Since MCI reordered the data during the ETL phase, previous reports and queries that were either very difficult to script or simply not possible are now done with relative ease. A graphical representation of the initial version of the new data warehouse database is shown below in Figure 12.

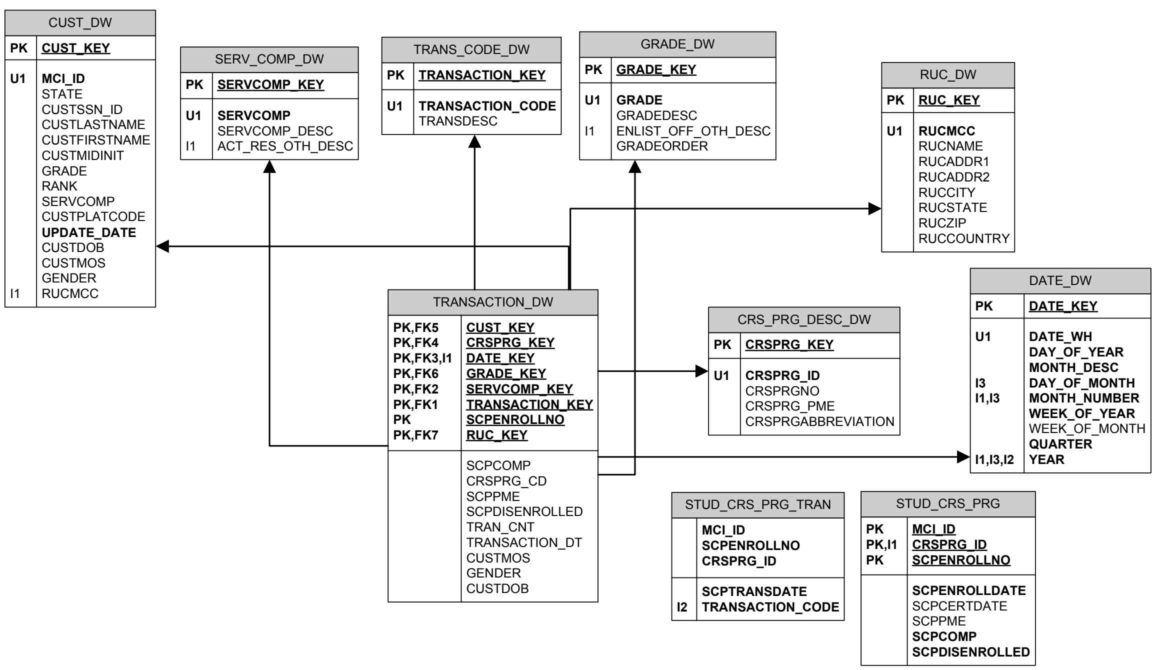


Figure 12. Data Warehouse Diagram

1. Description of the Warehouse Data Source

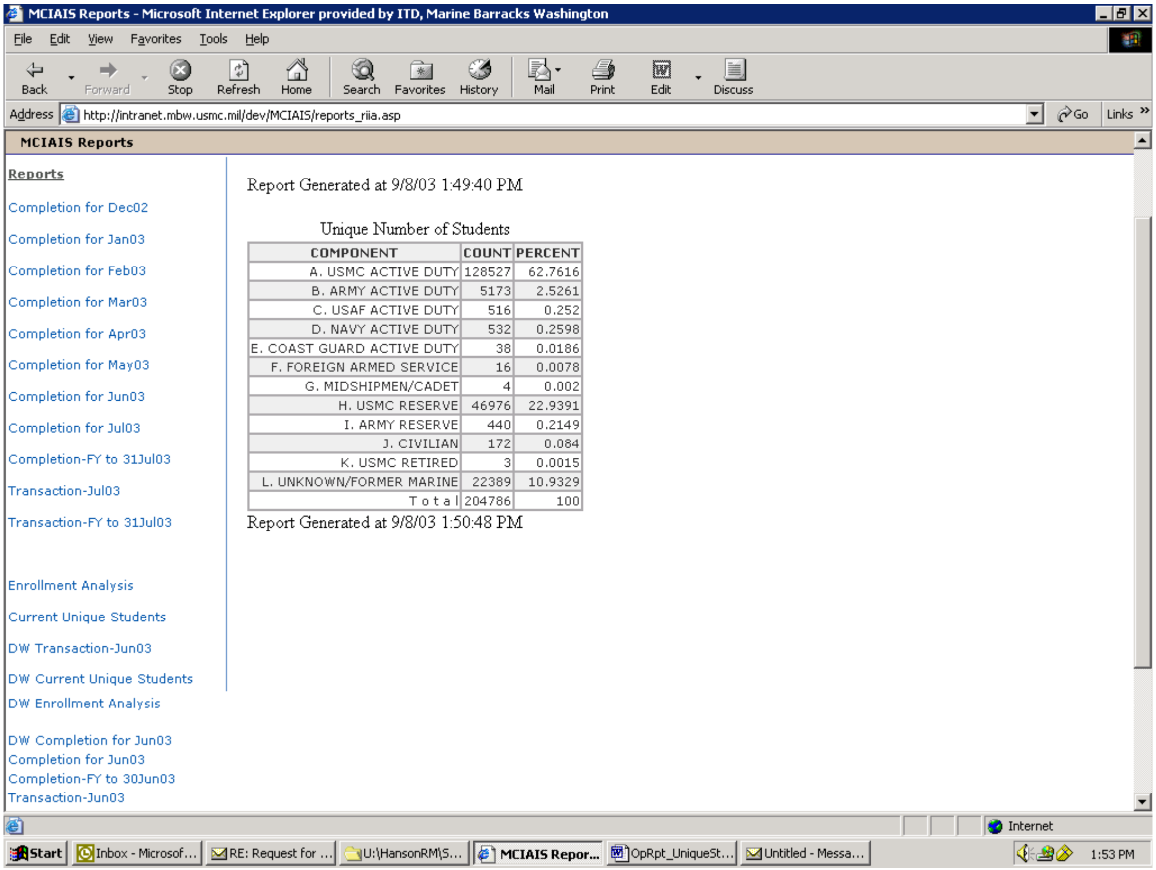
The diagram of the new database is much different than that described in detail in Chapter V. The 90+ tables contained in MCIAIS have been reduced to ten simpler, easier to use tables. As the functionality of the warehouse grows, more tables will undoubtedly be added, but nowhere near the ninety contained in the original.

Key elements contained in the new data source focus around the some of the same areas as broken down in Chapter V. They include the student, who is represented above in tables containing customer information, grade, service component, and RUC or unit. The other element deals with courses and transactions. These tables are course program, transaction, and transaction code. The linking tables that also are in the database are student course progress, student course transaction progress, and date to keep track of the date. Although this arrangement of data seems rather simple, it contains all of the key data required with which to replicate all of the queries and reports that were described in detail above.

2. Functionality

As previously stated, this new warehouse environment is much easier to query due to its simplicity. With some basic training, non-IT users will be able to query this data source themselves, freeing IT to press forward with the warehouse's untapped potential.

Phase I had the mandate of replicating the reports described above. Below are some of the new tables (Tables 5 and 6) pulled from the warehouse. They were designed to have the same look and feel of the previously generated reports, without losing any of the functionality.



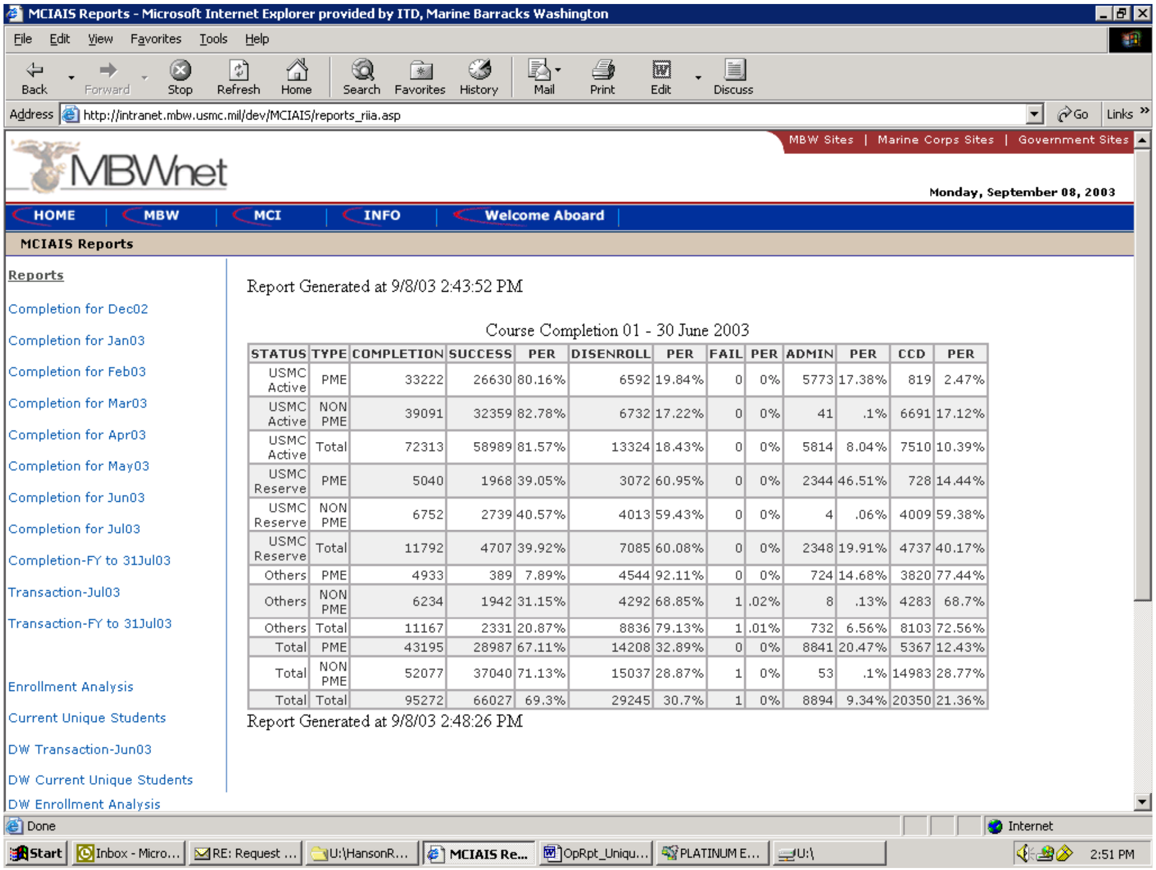
Report Generated at 9/8/03 1:49:40 PM

Unique Number of Students

COMPONENT	COUNT	PERCENT
A. USMC ACTIVE DUTY	128527	62.7616
B. ARMY ACTIVE DUTY	5173	2.5261
C. USAF ACTIVE DUTY	516	0.252
D. NAVY ACTIVE DUTY	532	0.2598
E. COAST GUARD ACTIVE DUTY	38	0.0186
F. FOREIGN ARMED SERVICE	16	0.0078
G. MIDSHIPMEN/CADET	4	0.002
H. USMC RESERVE	46976	22.9391
I. ARMY RESERVE	440	0.2149
J. CIVILIAN	172	0.084
K. USMC RETIRED	3	0.0015
L. UNKNOWN/FORMER MARINE	22389	10.9329
Total	204786	100

Report Generated at 9/8/03 1:50:48 PM

Table 5 Unique Students (Data Warehouse)



Report Generated at 9/8/03 2:43:52 PM

Course Completion 01 - 30 June 2003

STATUS	TYPE	COMPLETION	SUCCESS	PER	DISENROLL	PER	FAIL	PER	ADMIN	PER	CCD	PER
USMC Active	PME	33222	26630	80.16%	6592	19.84%	0	0%	5773	17.38%	819	2.47%
USMC Active	NON PME	39091	32359	82.78%	6732	17.22%	0	0%	41	.1%	6691	17.12%
USMC Active	Total	72313	58989	81.57%	13324	18.43%	0	0%	5814	8.04%	7510	10.39%
USMC Reserve	PME	5040	1968	39.05%	3072	60.95%	0	0%	2344	46.51%	728	14.44%
USMC Reserve	NON PME	6752	2739	40.57%	4013	59.43%	0	0%	4	.06%	4009	59.38%
USMC Reserve	Total	11792	4707	39.92%	7085	60.08%	0	0%	2348	19.91%	4737	40.17%
Others	PME	4933	389	7.89%	4544	92.11%	0	0%	724	14.68%	3820	77.44%
Others	NON PME	6234	1942	31.15%	4292	68.85%	1	.02%	8	.13%	4283	68.7%
Others	Total	11167	2331	20.87%	8836	79.13%	1	.01%	732	6.56%	8103	72.56%
Total	PME	43195	28987	67.11%	14208	32.89%	0	0%	8841	20.47%	5367	12.43%
Total	NON PME	52077	37040	71.13%	15037	28.87%	1	0%	53	.1%	14983	28.77%
Total	Total	95272	66027	69.3%	29245	30.7%	1	0%	8894	9.34%	20350	21.36%

Report Generated at 9/8/03 2:48:26 PM

Table 6 Course Completions (Data Warehouse)

C. USING THE CAPABILITIES OF THE WAREHOUSE (PHASE II)

As of the writing of this report, the warehouse has completed most of Phase I and has made some limited progress into Phase II. The hardware, software, and extraction, transfer, and load were completed, and several reports were moved on to the warehouse platform.

In depth discussions with key personnel at MCI have revealed where they are going to proceed, and based off of this, projections will now be put forward on what will be achieved.

The current management reports deal with courses and students. As laid out in earlier chapters, MCI is much more than courses and students, although they are a large part of the picture. MCIAIS has large subsets of data dealing with examinations, course ordering, logistics, and the maintenance and development of courses that are not adequately addressed in the current MCIAIS reports, or in their replicated warehouse equivalents. MCI will sit down with the "knowledge workers" in these critical areas and determine what questions will be of value for analysis, extract the appropriate data from MCIAIS, and transform and load it to the warehouse. Once there, it can be queried, analyzed and mined to the limits of the Oracle 9i system.

VII. KNOWLEDGE VALUE ADDED

As stated earlier in the review of measuring methodologies, KVA allows description of all process outputs (including those generated from IT) in common units. This allows allocation of revenue to IT in proportion to contributions to process outputs at the sub-corporate level. What this does, is show in measurable terms what value is actually added by implementing specific IT applications. Additionally, KVA allocates revenue as well as cost allowing for ratio analysis of the IT value-added.

It must be stressed that it is usual for KVA analysis to be conducted in conjunction with Business Process Reengineering. This is where processes are examined and altered to streamline the implementation of the IT solution. In the case of MCI's warehouse implementation, no process reengineering took place. The results of the return on knowledge for the three examples examined below can be completely attributed to the use of the capabilities of the warehouse. Should MCI decide to reengineer its processes to achieve further gain, that is their decision to make.

What follows is a measurement of how much value is actually added to the enterprise of MCI by the implementation of the warehouse.

A. THE MATHEMATICS BEHIND KVA

The individual analyses that are used in this thesis follow the formula described in Dr. Thomas Housel and Dr. Arthur Bell's book, Measuring and Managing Knowledge.

1. Knowledge Value Added Example

The formula has been entered into a Microsoft Excel spreadsheet for ease of calculation. The headings of the columns of a sample spreadsheet is shown below as Figure 13, and each column is described so the concept of the calculations are more easily understood.

Knowledge Value Added Example Spreadsheet

1	2		3	4	5	6	7	8		9	10
Process Elements	Rank	Act LT (days)	Rel LT	No. Emp	% IT	Knowledge in Automation (3*4*5)	Total Knowledge (3*4+6)	% Allocation of Knowledge	Annual Revenue Allocation (in US$)	Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	ROK
Write Paper	2	10	0.77	1	20%	0.15	0.92	76.92%	$7,692	$40,000	19.2%
Correct Paper	1	3	0.23	1	20%	0.05	0.28	23.08%	$2,308	$48,000	4.8%
Total	3	13	1.00				1.20	100.00%	$10,000	$88,000	**11.4%**
Correlation of 2 and 3	1										

Assumptions:	Annual Salary	$50,000
	Annual Salary	$60,000
	Target Cost Savings = "Revenue"	$10,000

Figure 13. Knowledge Value Added Spreadsheet Example

a. Process Elements

The Process Elements column is where the title of each step in the process to be analyzed is placed. In the example above, the sample process consists of writing and correcting a paper. These two steps are listed in column 1.

b. Rank and Actual Learning Time

In the Rank column, a number is placed indicating how difficult each step is in relation to the other steps. The more difficult a task is, the higher the number that is entered. The Actual Learning Time, in this case in units of days spent to learn the task is listed in the next column. A correlation of these is shown in the row on the bottom of column 1. In a perfect example, this correlation would be a value of 1. This is achieved in this example, as it takes longer to learn to write than to correct. In most actual implementations, this is not necessarily the case, but as long as the correlation is relatively high, the results will be valid.

c. Relative Learning Time

The Relative Learning Time is the learning time of the individual task as a percentage of the whole. This will be used in several of the follow-on calculations.

d. Number of Employees

The Number of Employees column shows how many individuals are required to perform this specific task. In the example, both tasks are performed by only one worker.

e. Percent IT

This column represents the amount of the process that is automated. This is an estimate, as it is very difficult to truly measure percentage of automation unless a process is completely done without the aid of computers (0%) or is completely automated (100%). In the example, 20% of both writing and correcting the paper are automated.

f. Knowledge in Automation

Column 6 is the product of columns 3, 4 and 5 (3*4*5). By multiplying relative learning time by number of employees performing the task, and then multiplying that result by the percentage of the process being automated, a number is produced representing the knowledge in automation.

g. Total Knowledge

Total Knowledge is generated by multiplying learning time by the number of employees and then adding the knowledge in automation to the result (3*4 + 5). This number is used in the next calculation, Percent Allocation of Knowledge. The Total Knowledge is the total value surrogate of the annual revenue.

h. Percent Allocation of Knowledge and Annual Revenue Allocation

Percent Allocation of Knowledge is the Total Knowledge for a step divided by the total contained in the column Total Knowledge, or the total knowledge in the process by task. Annual Revenue Allocation is the generated revenue (represented in the bottom box under assumptions), multiplied by the Percent Allocation of Knowledge.

The generated revenue is not always a specific amount, it is often a target, in the example case a desired savings of $10,000 is desired, so that is what is used in the calculations.

i. Annual Expense

The Annual Expense is the annual salary of the worker(s) multiplied by the percentage of knowledge NOT automated.

j. Return on Knowledge

The Return on Knowledge is simply revenue over expense. The total for the process as a whole is total revenue over total expense.

2. Results

What is obtained from entering values into the spreadsheet is a Return on Knowledge value. It is a percentage of the revenue in terms of knowledge over the expense in terms of knowledge of non-automated processes. This methodology allows the user to ascertain in terms of common units how much knowledge is actually contained in a process due to IT. At that point, additional IT is then embedded in the process. By comparing the Return on Knowledge of the original process or the "as-is" to the improved use of IT process or the "to-be", a way to measure the value of the new IT improvement is established.

Using this methodology, some specific MCI processes will be looked at viewing the current process in place, or the "as-is", and then comparing it to the process that will be in place embedding more IT using the data warehouse, or the "to-be".

B. INDIVIDUAL EXAMPLE ANALYSIS

To do a KVA analysis, existing processes that can be streamlined by use of the warehouse must be examined and measured in detail. The three processes chosen are querying and reporting from the data source, development of the annual plan, and course development. These processes as they currently exist are described in detail in Chapter IV. A brief recap of the process as it exists, and what elements will be affected by use of the warehouse accompany each analysis.

It must be noted that MCI is a non-profit organization, and thus the revenue is very difficult to figure out. In the cases measured below, the revenue figure is the $207,000 that it cost MCI to build the warehouse. What this is trying to show is how the cost of the warehouse will actually be made up using a better IT tool. Since the figure remains constant throughout the analyses it serves as a good reference that keeps all things equal from the to-be to the as-is in each case.

The first process examined is in the development and processing of the reports and queries themselves. Although looking at progress in an IT process using an IT tool seems a bit incestuous, it is a good example in this case, as it shows the considerable value already in the system, but also draws out the considerable additional benefits inherent in the warehouse.

1. Value-Added in Queries and Reports

In the current process using MCIAIS, any new report or query needs to be developed by the MCI user and one of the IT professionals on staff. This relationship must exist due to the fact that the MCI user knows what he wants to ask the database, and the IT professional has to write the SQL query to extract what is requested. That is if that specific query can be written at all. Once the question is finalized, then the IT professional writes the query. This requires a large volume of knowledge not only with SQL, but with the giant MCIAIS monster as it sits. The database is then queried by IT, and the results are formatted and made available on the MBW Intranet.

In the data warehouse environment, the process is very similar, although some key changes are made. The question must still be asked by the user, and at this point, they must still be assisted by an IT professional. The IT professional then writes the query, but uses a much simpler querying technique inherent in the warehouse itself. This step requires much less learning time as compared to teaching SQL and the background of MCIAIS. The query is then run, but is done so in a very automated format.

These critical variable changes have knowledge attached to them. KVA allows us to unlock just how much value. Contained below are spreadsheets with both processes, the old MCIAIS "as is" process, and the new warehouse "to-be" process. These are shown as Figures 14 and 15.

Query of MCIAIS "as-is"

1 Process Elements	2 Rank	Act LT (days)	3 Rel LT	4 No. Emp	5 % IT	6 Knowledge in Automation (3*4*5)	7 Total Knowledge (3*4+6)	8 % Allocation of Knowledge	Annual Revenue Allocation (in US$)	9 Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	10 ROK
Gather Input	3	10	0.25	2	20%	0.10	0.60	31.58%	$65,368	$98,896	66.1%
Write Query	4	20	0.50	1	70%	0.35	0.85	44.74%	$92,605	$18,543	499.4%
Query DB	2	5	0.13	1	90%	0.11	0.24	12.50%	$25,875	$6,181	418.6%
Format Output	1	5	0.13	1	70%	0.09	0.21	11.18%	$23,151	$18,543	124.9%
Total	10	40	1.00				1.90	100.00%	$207,000	$142,163	**145.6%**
Correlation of 2 and 3	0.912870929										

Assumptions: IT/DLTD Annual Salary	$61,810
IT Annual Salary	$61,810
IT EA Salaries	$61,810
Supervisor Salary	$80,000
Target Cost Savings = "Revenue"	$207,000

Figure 14. Query of MCIAIS "as-is"

Query of Data Warehouse "to-be"

1 Process Elements	2 Rank	Act LT (days)	3 Rel LT	4 No. Emp	5 % IT	6 Knowledge in Automation (3*5)	7 Total Knowledge (3*4+6)	8 % Allocation of Knowledge	Annual Revenue Allocation (in US$)	9 Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	10 ROK
Gather Input	3	10	0.25	2	70%	0.35	0.85	37.18%	$76,960	$37,086	207.5%
Write Query	4	20	0.50	1	90%	0.45	0.95	41.55%	$86,014	$6,181	1391.6%
Query DB	2	5	0.13	1	99%	0.12	0.25	10.88%	$22,522	$618	3643.8%
Format Output	1	5	0.13	1	90%	0.11	0.24	10.39%	$21,504	$6,181	347.9%
Total	10	40	1.00				2.29	100.00%	$207,000	$50,066	
Correlation of 2 and 3	0.912870929										

Assumptions: IT/DLTD Worker Salary	$61,810
IT Annual Salary	$61,810
IT Salaries	$61,810
Supervisor Salary	$80,000
Target Cost Savings = "Revenue"	$207,000

Figure 15. Query of Data Warehouse "to-be"

The critical differences in the example the amount of IT contained in the process steps. Since the data warehouse is much more user friendly and doesn't require an extensive knowledge of SQL, the knowledge required for the step "write query" is much more significantly embedded in IT. Additionally, querying the database and formatting output are now significantly easier to learn and have much more automation in their processes. By changing these variables, the return on knowledge rose from a respectable 146.8% to a much higher 418.6%.

In this formula, the cost of the warehouse is leveraged with the cost of the individuals running the processes. What results is a net gain of 271.8%. In many KVA examples, the "as-is" process has a very low percentage in return on knowledge, due to a small amount of IT inherent in the old process. In this case, querying MCIAIS has a good level of IT in the old way, so the number is fairly good at over 100%. By using the IT tool of the warehouse to further the capability, the vastly improved results are achieved.

2. Value Added in Course Development

In Chapter IV, the current eight-step process for course development is outlined in detail. There are several steps in this process using embedded knowledge in the warehouse that will show value-added.

The process as outlined has the steps of Initial Planning Conference, Analysis, Design, Pre-Development Conference, Development, Implementation, Evaluation Planning Conference, and Evaluation. Although the warehouse has abundant capabilities, in this case, the results achieved are not that dramatic. Preparation for the conferences can have added IT imbedded through use of the warehouse. Additionally, analysis and design are likewise affected by using the capabilities of the warehouse to do more of the task more efficiently through maximizing IT. The analysis of the Knowledge Value Added is shown below as Figures 16 and 17.

Course Development "as-is"

1	2		3	4	5	6	7	8		9	10
Process Elements	Rank	Act LT (days)	Rel LT	No. Emp	%IT	Knowledge in Automation (3*4*5)	Total Knowledge (3*4+6)	% Allocation of Knowledge	Annual Revenue Allocation (in US$)	Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	ROK
Initial Planning Conference	3	3	0.03	6	10%	0.02	0.20	3.95%	$8,180	$333,774	2.5%
Analysis	7	20	0.20	3	15%	0.09	0.70	13.77%	$28,506	$157,616	18.1%
Design	6	20	0.20	4	30%	0.24	1.05	20.76%	$42,966	$173,068	24.8%
Pre-Development Conference	2	3	0.03	6	10%	0.02	0.20	3.95%	$8,180	$263,671	3.1%
Development	8	25	0.25	3	40%	0.30	1.06	20.96%	$43,379	$66,226	65.5%
Implementation	5	15	0.15	5	25%	0.19	0.95	18.71%	$38,731	$231,788	16.7%
Evaluation Planning Conference	1	3	0.03	6	15%	0.03	0.21	4.13%	$8,552	$315,231	2.7%
Evaluation	4	10	0.10	6	15%	0.09	0.70	13.77%	$28,506	$315,231	9.0%
Total	36	99	1.00				5.06	100.00%	$207,000	$1,856,604	
Correlation of 2 and 3	0.968871645										

Assumptions: IT/DLTD Annual Salary	$61,810
Program Manager Annual Salary	$48,828
Course Developer Annual Salary	$36,792
Supervisor Salary	$73,503
Target Cost Savings = "Revenue"	$207,000

Figure 16. Course Development “as-is”

Although not as dramatic as the first example, these results are encouraging. A process that is not primarily designed to be a beneficiary of the warehouse has managed to double its return on knowledge from a paltry 11.1% to a somewhat improved 23.0%. This process is a prime candidate for a complete business process reengineering analysis to improve the return on knowledge. The warehouse can help in its small way, but the underlying problems must be solved first.

Course Development "to-be"

1	2		3	4	5	6	7	8		9	10
Process Elements	Rank	Act LT (days)	Rel LT	No. Emp	% IT	Knowledge in Automation (3*4*5)	Total Knowledge (3*4+6)	% Allocation of Knowledge	Annual Revenue Allocation (in US$)	Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	ROK
Initial Planning Conference	3	3	0.03	6	65%	0.13	0.33	5.44%	$11,255	$129,801	8.7%
Analysis	7	20	0.22	3	70%	0.46	1.12	18.67%	$38,653	$55,629	69.5%
Design	6	20	0.22	4	70%	0.62	1.49	24.90%	$51,537	$74,172	69.5%
Pre-Development Conference	2	3	0.03	6	65%	0.13	0.33	5.44%	$11,255	$102,539	11.0%
Development	8	25	0.27	3	40%	0.33	1.15	19.22%	$39,789	$66,226	60.1%
Implementation	5	15	0.16	5	25%	0.21	1.03	17.16%	$35,526	$231,788	15.3%
Evaluation Planning Conference	1	3	0.03	6	65%	0.13	0.33	5.44%	$11,255	$129,801	8.7%
Evaluation	4	2	0.02	6	70%	0.09	0.22	3.73%	$7,731	$111,258	6.9%
Total	36	91	1.00				6.00	100.00%	$207,000	$901,213	23.0%
Correlation of 2 and 3	0.968872										

Assumptions: T/DLTD Annual Salary	$61,810
Program Manager Annual Salary	$48,828
Course Developer Annual Salary	$36,792
Supervisor Salary	$73,503
Target Cost Savings = "Revenue"	$207,000

Figure 17. Course Development “to-be”

3. Value Added in Strategic Decision Making

The strategic decision making process at MCI is where all of the executive-level individuals determine the course that MCI will take in any given year. Factors that must be weighed include analysis of previous years plans, analysis of new requirements, manpower available, and the budget under which these tasks are to be accomplished.

All of these factors determine what will be produced and what will fall by the wayside. The data warehouse analysis capability of analyzing requirements and manpower data, and the previous years plans will cut down on the learning time required to perform these functions and increase the percentage of IT in those analyses. Returns on knowledge are illustrated below in Figures 18 and 19.

Annual Plan "as-is"

1	2		3	4	5	6	7	8		9	10
Process Elements	Rank	Act LT (days)	Rel LT	No. Emp	% IT	Knowledge in Automation (3*4*5)	Total Knowledge (3*4+6)	% Allocation of Knowledge	Annual Revenue Allocation (in US$)	Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	ROK
Strategic Vision Conference	2	3	0.03	5	10%	0.01	0.15	3.77%	$7,798	$323,514	2.4%
Previous Plan Analysis	8	21	0.20	3	30%	0.18	0.77	18.70%	$38,706	$154,356	25.1%
New Requirements Analysis	7	21	0.20	3	30%	0.18	0.77	18.70%	$38,706	$154,356	25.1%
Manpower Analysis	6	21	0.20	3	30%	0.18	0.77	18.70%	$38,706	$154,356	25.1%
Fiscal Analysis	5	21	0.20	3	40%	0.24	0.82	20.14%	$41,684	$132,305	31.5%
Analysis Conference	3	3	0.03	5	10%	0.01	0.15	3.77%	$7,798	$459,837	1.7%
In-Process Revision	4	14	0.13	3	30%	0.12	0.51	12.47%	$25,804	$154,356	16.7%
Decision Conference	1	3	0.03	5	10%	0.01	0.15	3.77%	$7,798	$459,837	1.7%
Total	36	107	1.00				4.09	100.00%	$207,000	$1,992,919	**10.4%**
Correlation of 2 and 3	0.881090309										

Assumptions:		
	Department Chief Salary	$73,503
	Deputy Director Annual Salary	$71,892
	Executive Director Annual Salary	$102,186
	IT Professional Salary	$61,810
	Target Cost Savings = "Revenue"	$207,000

Figure 18. Annual Plan "as-is"

The returns on knowledge in both of the cases is very low, not due to poor learning times or poor use of IT, but due to the high salaries of the workers involved in the processes. These analyses and decision conferences are being held by senior civilian government service workers (GS 13-15), and military officers holding the ranks of Major and Lieutenant Colonel. Because of this the expense of the tasks is very high, and the value of the denominator of the equation is very high. The ROK is maintained at a low level because of this.

Annual Plan "to-be"

1	2		3	4	5	6	7	8		9	10
Process Elements	Rank	Act LT (days)	Rel LT	No. Emp	% IT	Knowledge in Automation (3*4*5)	Total Knowledge (3*4+6)	% Allocation of Knowledge	Annual Revenue Allocation (in US$)	Annual Expense (Salary X No. Emp X %Knowledge NOT in IT	ROK
Strategic Vision Conference	2	3	0.03	5	70%	0.10	0.24	4.39%	$9,093	$107,838	8.4%
Previous Plan Analysis	8	21	0.20	3	80%	0.47	1.06	19.53%	$40,437	$44,102	91.7%
New Requirements Analysis	7	21	0.20	3	80%	0.47	1.06	19.53%	$40,437	$44,102	91.7%
Manpower Analysis	6	21	0.20	3	80%	0.47	1.06	19.53%	$40,437	$44,102	91.7%
Fiscal Analysis	5	21	0.20	3	40%	0.24	0.82	15.19%	$31,451	$132,305	23.8%
Analysis Conference	3	3	0.03	5	70%	0.10	0.24	4.39%	$9,093	$153,279	5.9%
In-Process Revision	4	14	0.13	3	80%	0.31	0.71	13.02%	$26,958	$44,102	61.1%
Decision Conference	1	3	0.03	5	70%	0.10	0.24	4.39%	$9,093	$153,279	5.9%
Total	36	107	1.00				5.43	100.00%	$207,000	$723,109	28.6%
Correlation of 2 and 3	0.881090309										

Assumptions:		
	Department Chief Salary	$73,503
	Deputy Director Annual Salary	$71,892
	Executive Director Annual Salary	$102,186
	IT Professional Salary	$61,810
	Target Cost Savings = "Revenue"	$207,000

Figure 19. Annual Plan "to-be"

A comparison between the two cases reveals values of ROK of 10.4% for the "as-is" and 28.6% for the "to-be". This is again a doubling of knowledge value added. To truly being these numbers up, the work performed needs to be delegated to less salaried workers, but due to the nature of these decisions, this is very unlikely to happen.

C. AGGREGATE KNOWLEDGE VALUE ADDED

By looking at these three examples, in each case the return on knowledge was over double in the "to-be" as in the "as-is", with the IT example being significantly greater. These results only scratch the surface of the capabilities of the warehouse, and its value to MCI and its core processes. Once Student Service Department's logistics section, and more of the strategic analysis by the headquarters begin to adapt to the tool, even more value will be extracted from the warehouse implementation.

VIII. CONCLUSION

The purpose of this thesis was to evaluate the implementation of a data warehouse in the specific military environment of the Marine Corps Institute. The measuring tool used for this evaluation was Knowledge Value Added.

Through the review of capabilities of data warehouses and measuring methodologies, it was shown that data warehouses can extend the IT capabilities of an enterprise such as MCI, and the amount of knowledge value added can be properly measured using KVA.

MCI is a complicated entity that produces and administers distance-learning materials for students worldwide. To accomplish this task, they have a well thought out organizational structure, and core processes that allow them to fulfill their mission in a successful manner. To support their endeavors, a large database has evolved to keep track of their processes. In addition to this, this database serves as the data source for their connection to the web, as well as an interface to the Marine Corps' main administrative database.

The data warehouse implementation was envisioned to allow for more timely completion of the current querying and reporting done on MCIAIS, and to open up all of the new warehouse capabilities such as OLAP, statistical analysis and data mining. Additionally, the data warehouse was to streamline current processes and provide easier access for the user to refined data in the completion of their mission.

KVA analysis has shown that the implementation has been worth the effort and resources expended. For a very low cost of $207,000, return on knowledge has at least doubled, and in one case increased by a factor of four. These figures do not include the additional benefits that will be reaped when the warehouse is completely deployed and all of the OLAP, statistical analysis, and data mining tools are operational.

As with all IT implementations, it is imperative that MCI continue to use the data warehouse to its fullest capabilities. A very powerful and valuable tool is in their hands that has been shown to double their return on knowledge. That can only continue if they continue to exploit its capabilities.

These gains should not be MCI's alone. There are many enterprises within the Department of Defense that have built up large amounts of data in numerous data sources. By using already purchased site licenses from vendors such as Oracle, and buying relatively inexpensive hardware, institutions can reap the same benefits as MCI through the implementation of data warehousing solutions.

IX. RECOMMENDATIONS

The progress made by MCI to date is very encouraging. By spending a mere $207,000, a platform is in place that will allow a much higher return on knowledge through increased use of IT. The process must not end here, however.

A. EDUCATING THE WORKERS

Now that the IT department has the warehouse set up and able to generate queries and reports from the warehouse data source, the workers at MCI in the other departments must be made aware of the new capabilities that exist. They must be instructed on the capabilities and limitations of the warehouse, so when they wish to get answers from the data, they know the parameters that they must work within. Additionally, to get a better payoff on the knowledge value added of the warehouse, the workers need to learn how to use the warehouse's capability themselves. By self-generating the queries, which is now well within expectations due to the ease of use of the warehouse, the IT worker will no longer need to generate the query for the normal worker. This will eliminate the task from the IT workers list, and effectively make this a one-worker task vice two, increasing overall institute productivity.

B. FURTHER DEVELOP THE WAREHOUSE'S CAPABILITIES

As stated in the description of the warehouse's capabilities, the Oracle 9i data warehouse is capable of performing online analytical processing, statistical analysis, and data mining. At this point, MCI has only a replication of what it had from MCIAIS with better functionality. With the warehouse in place, questions can be asked of the data that in the old environment were simply not answerable. What has to happen now is for the questions to be asked. All of the non-IT departments need to do some introspection to determine what the data may be able to tell them. You can't get an answer if you don't ask the question.

Additionally, the data mining features need to be fully deployed. MCI may be sitting on the mother lode of information, but without panning for the nuggets, they will never know. By using the inherent functionality in the system, they now have the IT pick and shovel required for mining. Like the forty-niner of old, they just need to dig.

BIBLIOGRAPHY

1. Adelman, Sid, cited 2003, Ten Worst Practices of the Unsuccessful Data Warehouse Project Manager, available online at: [http://doe.k.12.hi.us/technology/businteI/sa_dw10worstpractices.htm] September 2003.

2. CEISS Research and IT Solutions, cited 2003, A Case Study of Data Warehousing, available online at: [http://www.ceiss.org/products/case-odu.asp]. September 2003.

3. El Sawy, Omar, A., 2001: *Redesigning Enterprise Processes for e-Business*, McGraw Hill, 206 pp.

4. Fingar, Peter, and Aronica, Ronald, 2001, *The Death of "e" and the Birth of the Real New Economy*, Meghan-Kiffler Press, 339 pp.

5. Greenfield, Larry, cited 2003, The Data Warehousing Information Center, available online at: [http://www.dwinfocenter.org/]. September 2003.

6. Housel, Thomas, J., et.al., December, 2002: "Requirements for a Theory of Measuring the Return on Information Technology: A Proof of Concept Demonstration.

7. Housel, Thomas J., and Bell, Arthur H., 2001: *Measuring and Managing Knowledge*, McGraw Hill, 162 pp.

8. Kimball, Ralph., 1996: *The Data Warehouse Toolkit*, John Wiley and Sons, Inc., 367 pp.

9. Multiple Authors, cited 2002-2003, Marine Corps Institute Online, available online at: [mci.usmc.mil]. September 2003.

10. Ohio State University, cited 2003, The Information Warehouse: Analysis at Your Fingertips, available online at: [medicine.osu.edu/Informatics/talks/powerpoint/CCG%20Presentation.ppt]. September 2003.

11. Oracle Corporation, cited 2003, Oracle 9i for Data Warehousing, and other topics, available online at [www.oracle.com]. September 2003.

12. Stair, Ralph M., and Reynolds, George W., 2001: *Principles of Information Systems*, Fifth Edition, Course Technology, 305 pp.

13. Wu, M.C., and Buchmann, A.P., cited 2003: Research Issues in Data Warehousing, available online at: [http://www.dvs1.informatik.tu-darmstadt.de/staff/wu/wu.htm]. September 2003.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Marine Corps Representative
   Naval Postgraduate School
   Monterey, California

4. Director, Training and Education, MCCDC, Code C46
   Quantico, Virginia

5. Director, Marine Corps Research Center, MCCDC, Code C40RC
   Quantico, Virginia

6. Marine Corps Tactical Systems Support Activity (Attn: Operations Officer)
   Camp Pendleton, California

7. Executive Director, Marine Corps Institute
   Marine Barracks, Washington, DC

8. Chair, Information Sciences Department
   Naval Postgraduate School
   Monterey, California

9. Dr. Thomas J. Housel
   Naval Postgraduate School
   Monterey, California

10. Glenn R. Cook
    Naval Postgraduate School
    Monterey, California