

Data Economics in Army Information Systems

AD-A211 582

This document has been approved
for public release and sale by
the MITRE Corporation.

OTIC
82

89

8 21 029

MITRE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release, distribution unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE					
4. PERFORMING ORGANIZATION REPORT NUMBER(S) MTR-89W00029			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION The MITRE Corporation	6b. OFFICE SYMBOL (if applicable) W113	7a. NAME OF MONITORING ORGANIZATION U.S. Army Information System Command (ASPL-AD)			
6c. ADDRESS (City, State, and ZIP Code) 7525 Colshire Drive McLean, VA 22102		7b. ADDRESS (City, State, and ZIP Code) Fort Huachuca, AZ 85613			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F19628-89-C-0001			
8c. ADDRESS (City, State, and ZIP Code)		10. SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Data Economics in Army Information Systems					
12. PERSONAL AUTHOR(S) Coonce, Thomas J. and Coutavas, Spiros D.					
13a. TYPE OF REPORT Final	13b. TIME COVERED FROM 12/88 TO 2/89	14. DATE OF REPORT (Year, Month, Day) 89/03	15. PAGE COUNT 176		
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
05	02		Army information systems, data quality, data standardization,		
12	07		information value, MIS evaluation.		
19. ABSTRACT (Continue on reverse if necessary and identify by block number) The Army stores and maintains redundant data within and between Army information systems. Moreover, many of these data are not standardized with respect to definition, composition, currency, or value. This situation may give rise to poor decision-making by users of these systems. Faulty decisions may lead to improper allocation of resources, and result in cost, schedule, and technical performance penalties. This MITRE Technical Report (MTR) addresses the hypothesis that inconsistent, non-standardized data, in addition to their inherent cost inefficiencies, can lead to faulty decision-making with quantifiable impacts on Army resource allocations. This study is based on a review of existing research and discussion with information specialists. It summarizes the cost of data redundancy and identifies promising cost-benefit and information economics methodologies that provide useful tools for linking data quality to the quality of decisions. A synthesized evaluation framework is derived for further testing and application within a representative Army personnel data sharing environment. <i>Recommends</i>					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL Spiros D. Coutavas			22b. TELEPHONE (Include Area Code) (703) 883-7752	22c. OFFICE SYMBOL W113	

Data Economics in Army Information Systems

Thomas J. Coonce
Spiros D. Coutavas

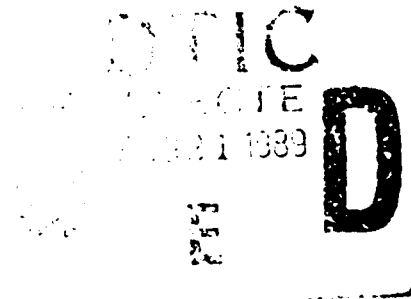
March 1989

MTR-89W00029

SPONSOR:
USAISC
CONTRACT NO.:
F19628-89-C-0001

It has been approved for public release.

The MITRE Corporation
Washington C3I Division
7525 Colshire Drive
McLean, Virginia 22102-3481



MITRE Project Approval:

A handwritten signature in dark ink, appearing to read "W. P. Hutzler", written over a horizontal line.

Technical Center Approval:

A handwritten signature in dark ink, appearing to read "William P. Hutzler", written over a horizontal line.

William P. Hutzler Manager
Economic Analysis Center

ABSTRACT

The Army stores and maintains redundant data within and between Army information systems. Moreover, many of these data are not standardized with respect to definition, composition, currency, or value. This situation may give rise to poor decision-making by users of these systems. Faulty decisions may lead to improper allocation of resources, and result in cost, schedule, and technical performance penalties. This MITRE Technical Report (MTR) addresses the hypothesis that inconsistent, non-standardized data, in addition to their inherent cost inefficiencies, can lead to faulty decision-making with quantifiable impacts on Army resource allocations. This study is based on a review of existing research and discussions with information specialists. It summarizes the cost of data redundancy and identifies promising cost-benefit and information economics methodologies that provide useful tools for linking data quality to the quality of decisions. A synthesized evaluation framework is derived for further testing and application within a representative Army personnel data sharing environment.

This study report was prepared for the Command and Control Information Systems Department, W-114, Command Center Technology Division, by the MITRE Washington Economic Analysis Center (EAC).

Suggested keywords: Army information systems, data quality, data standardization, information value, MIS evaluation.



Accession For

NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	

Date _____

A-1

TABLE OF CONTENTS

SECTION	PAGE
List of Figures	vi
List of Tables	vii
Executive Summary	ix
1 Introduction	1-1
1.1 Background	1-1
1.2 Purpose of Research Effort	1-2
1.3 Research Approach	1-3
2 Findings and Analysis	2-1
2.1 Research Framework	2-1
2.1.1 Problem Statement	2-1
2.1.2 Research Questions	2-2
2.1.3 Definitions of Key Terms	2-2
2.2 The Costs and Benefits of Data	2-6
2.2.1 The Cost of Data Redundancy in Single Systems	2-6
2.2.2 The Costs and Benefits of a Shared Data Environment	2-8
2.2.3 The Value of Data and Information	2-13
2.2.4 A Conceptual Framework for Information System Cost and Benefit Evaluation	2-28
3 Conclusions and Recommendations	3-1
3.1 Survey of Previous Research	3-1
3.2 Assessment of Data Redundancy	3-2
3.3 Evaluation of Approaches for Measuring Impact of Data Quality on Information Systems	3-4
3.4 Recommendations	3-4
APPENDIX A: Document Abstracts	A-1
APPENDIX B: Document Summaries	B-1
REFERENCES	R-1
GLOSSARY	G-1

LIST OF FIGURES

FIGURE	PAGE
2-1 Horizontal and Vertical Integration	2-4
2-2 Levels of Technical Integration	2-5
2-3 Payoff Versus Data Quality Improvement Cost	2-22
2-4 Hypothetical Army Data Delivery Systems for Three Functional Areas	2-29

LIST OF TABLES

TABLE	PAGE
ES-1 Overview of Selected Approaches for Evaluation of Information Systems	xiv
ES-2 Proposed Report Format of NPV Costs and Benefits for Army Stove-Pipe Versus Shared Data Environments	xvii
2-1 Results of TAPDB Cost Benefit Analysis Intercomponent Data Access Alternatives	2-12
2-2 Overview of Selected Approaches for Evaluation of Information Systems	2-21
2-3 Proposed Report Format of NPV Costs and Benefits for Army Stove-Pipe Versus Shared Data Environments	2-32

EXECUTIVE SUMMARY

INTRODUCTION

One of the major issues facing the Army today is the need to share data from information systems among functional components in an automated fashion. In the past, Army applications were developed independently to support only the functional proponent. Little regard was given to the support these information systems could give other functions. This situation resulted in redundant data files, inconsistent data and inaccessibility of data from the many operational applications in the various Army functions. Data were typically locked into a proponent's application which had limited data compatibility and inhibited the Army's ability to share data. The inability to share data or operate in a more standardized data environment has broad implications, not only for information system efficiency, but also for the Army's reliance on these systems in its decision-making activities.

This report summarizes previous research concerning the costs and benefits of shared data environments and non-shared data environments in both military and commercial settings. The report also analyzes the pros and cons of methodologies applied to measure costs and benefits of information systems in general, and processed data, in particular, and recommends a specific approach to be applied to an Army data sharing situation.

PURPOSE AND SCOPE

The primary purpose of this report is to summarize research on the quantification of the benefits of a shared data environment (or reduced data redundancy) within either military or commercial information systems, and the effect on decision-making. This effort includes an assessment of the applicability of available evaluation methodologies to an Army data environment.

BACKGROUND

Since 1982, the Army has been focusing increasing attention on managing information as a resource. The Army has established an Information Mission Area (IMA) to ensure the integration of data resources. An organizational structure has been created to support this objective. The Office of the Directorate of Information Systems for Command, Control, Communications, and Computers (DISC4) is the Army Headquarters component responsible for oversight and policy development for the IMA. The Information Systems

Command (ISC) is responsible for the implementation of the IMA; and the Information System Engineering Command (ISEC), a subordinate command under ISC, has the responsibility for technical development of several IMA functions. Development of an Army Data Management Program and an Army Corporate Data Base (ACDB) were two of the major functions assigned to the ISEC.

The Army's basic strategy for meeting the objectives of the IMA is to formalize the process by which new systems are built, and to standardize data elements. This strategy is specified in Army Regulation (AR) 25-1, The Army Information Program, and the process to standardize data is guided by AR 25-9, The Army Data Management and Standards Program. These are the mechanisms by which the Army ensures data integration, sharing, standardization, interoperability, timeliness, and validity of information provided to Army decision-makers. The program provides Army leadership with a tool for identifying existing information resources, validating and satisfying known information requirements, and providing a systematic approach for acquiring future resources.

The ACDB concept was one alternative for meeting the objectives of the IMA. It was designed to further the aim of data sharing and to reduce redundant data caused by multiple data bases. The initiative was abandoned because of the lack of quantitative data to justify the projected expense. Many people in Army leadership intuitively "knew" that a shared data base environment was better than the many separate, disparate data base environments, but it seems that no one had specifically quantified the benefits. This is the major motivation for this research effort.

APPROACH

MITRE performed a detailed literature search of the military, academic, and commercial references. The primary focus of this effort was to determine if anyone had applied quantitative techniques toward measuring and comparing the effectiveness of information systems, including the benefits and costs of developing and operating a shared data environment versus a group of separate data bases. Each document related to the topic was reviewed and an abstract was written. MITRE reviewed over 100 documents and prepared abstracts for over 60. These abstracts are contained in Appendix A.

MITRE also obtained background data and collected applicable research documents by interviewing information specialists from the Department of Defense (DOD), academia, and the Management Information System (MIS) field. The content of these interviews are contained in Appendix B.

MITRE summarized and analyzed all documents reviewed. This information is contained in Section 2. Analysis consisted of evaluating the methodology used to address the cost and value of processed data in general, and the

cost and benefits of a shared data environment versus a non-shared data environment in particular. Any relevant methodologies discovered in the literature search were assessed for applicability to the Army environment. Conclusions and recommendations were then formulated; these are contained in Section 3 of the main body of the report.

SUMMARY OF FINDINGS

The Army's basic hypothesis is that "stove-pipe" information systems cause "redundant" and inconsistent data within the Army echelons and this causes inefficient and ineffective decision-making. Army leadership feels that the decision-making process and outcomes can be substantially improved if the Army moves toward integrating data among its functional components. MITRE's research efforts were focused on searching the literature to discover if anyone had measured the costs and benefits of a shared data environment versus a non-shared data environment.

Very little quantitative research has been done which definitively answers the question of whether a shared data environment is more cost-effective than non-shared data environments. Moreover, few documents were found which measured the value of processed data within a single information system. However, some research has been performed which grossly estimates the cost of "data redundancy" within a single information system. Other research measured the costs and benefits of an Army data sharing environment, and a number of researchers have written about the various methods of valuing processed data (information) from efficiency and effectiveness viewpoints. The salient findings from our literature search are summarized below.

1. Data redundancy--defined as a state in which data are defined and/or maintained by more than one source. Data redundancy is prevalent in many organizational information systems and is roughly estimated at 75 percent.
2. The cost of data redundancy within single corporation information systems has been roughly estimated at 50 percent of an organization's software maintenance effort.
3. Data redundancy and data inconsistency--defined as the same named data element appearing in two information systems which do not contain the same information or which possess different formats--are controllable with a strong Data Administrator, and forced use of automated data element dictionaries and Computer Aided Software Engineering (CASE) tools. Forced use of automated tools throughout the software life cycle creates data standardization and can substantially reduce software development and maintenance costs and moderate hardware requirements. Benefits of the overall

data administration have been estimated by several researchers who project savings of 10 to 15 percent in development and 80 percent in systems maintenance. However, no research was found which rigorously measured these potential savings.

4. Potential benefits of a shared data environment in the corporate setting were stated as:
 - Return on investment will be increased due to lower costs
 - Better information will permit the firm to attain a more competitive edge, maintain market leadership or enter a new product/market area
 - Software development cost will be shared
 - Maintenance costs will be shared
 - System reliability will be increased
 - Operating costs will be shared
 - Resource allocation decisions will be more efficient
5. One researcher found that a wholesale distributor achieved a 30 percent return on the investment in a shared data system, primarily because of a 40 percent reduction in order/buying personnel and a reduction in processing equipment costs.
6. In spite of such possible benefits, one researcher found that 25 percent of all real-time and telecommunications based software projects fail, primarily because of the "bad data" and the difficulty of getting the meaning of data straight among organizational users.
7. Within the Army environment, we found one study which measured the incremental costs and benefits of the various alternate ways to share personnel data among four personnel components. This analysis compared the costs and benefits of technically integrating four personnel systems over the baseline of merely sending data tapes. Results were presented in cost-to-benefit ratios and the best ratio was achieved by a remote access alternative. Benefits in this study were measured in data processing efficiency terms only. The value of obtaining better information for resulting decisions was not addressed or quantified.
8. Measuring the benefits of better data which can accrue from a data sharing environment centers around addressing the more basic

question of measuring the value of processed data (information) in and of itself. In particular, the question of whether data sharing is better than a non-shared environment boils down to addressing both the efficiency of the information delivery system and the effectiveness of the decision outcomes from "better" information.

9. Data on information value to the user is a critical concept, often overlooked, that is just now receiving the attention it demands. The notion that an information system should be justified on the basis of its usage rather than solely by its efficiency has too often been relegated to the "back burner" as "too hard". However, information specialists and academics have kept the issue alive with a considerable flow of "think piece" papers, theoretical approaches, conceptual frameworks, and occasional applications. The federal government, with the Paperwork Reduction Act of 1980 and the resulting Information Resources Management (IRM) concept has reinforced the characterization of information as a resource that needs to be managed, and has spurred further interest in information cost and value.
10. No operations research methodologies were found in the literature search that appeared feasible for application to the information valuation program. Although decision theory/game theory provide methodologies for incorporating many of the complexities in the quality of data versus quality of decision-making issue, the fact remains that the subjective probability distributions required by the analysis must be obtained by a degree of user interaction that is better accommodated in other analytical frameworks and approaches.
11. Mirroring the differences in philosophy and orientation between system developers and system users, two basic schools of thought on justifying investment in information technology systems have arisen. Developers key on the efficiency of the information process, i.e., the degree of cost-efficiency of combining resources to produce an information output, while user/decision-makers are primarily concerned with the effectiveness with which the output is used to enhance organizational/user performance. The resultant approaches arising from the efficiency and effectiveness orientations are cost-benefit analysis and decision theory, respectively. The approaches are narrowly defined to imply that they look exclusively at values either in information production or information use. The narrow focus corresponds closely to the manner in which these approaches are actually used in information system analysis. A third generic approach, which is referred to as Information Life Cycle Analysis (ILCA), combines elements of both the basic approaches and examines the costs and

TABLE ES-1
OVERVIEW OF SELECTED APPROACHES FOR EVALUATION
OF INFORMATION SYSTEMS

APPROACH	ASSUMPTIONS	OBJECTIVES	INFERENCES	FUNCTION VS PROCESS	STAGE OF LIFE CYCLE	INFORMATION TRANSFER VS USE	KEY INFORMATION VALUE CRITERIA
COST-BENEFIT ANALYSIS	Increased information needs are critical. Technological innova- tions can be used to address this challenge.	The development of efficient storage and retrieval systems.	Given efficient data management, use will automatically follow.	Process	Storage and Retrieval	Transfer	Technological factors-- reliability, simplicity, flexibility, adaptability.
INFORMATION LIFE CYCLE ANALYSIS	Concern with increasing data related costs. Quality, not quantity, is the critical issue.	Improve communi- cations between users and proces- sors. Improve decision-making frameworks.	Minimize redundancy and improve quality of information. More accurate and relevant data for more percep- tive decision-making.	Both	All	Both	Credibility of data-- accuracy, timeliness, and relevance. Value of assessments/ judgments made by users.
DECISION THEORY	Quality of information is important. Quantity and quality are both controlled--unavailable or too costly.	The optimal amount of information to make a decision.	Given a certain quanti- ty of information, a decision maker can reduce uncertainty and maximize his utility.	Function	Analysis and Use of information	Use	Relevance of data-- effectiveness in reducing uncertainty/ increasing utility. Timeliness, accuracy.

benefits over the entire information cycle (from production to use). The three approaches are summarized in table ES-1.

CONCLUSIONS AND RECOMMENDATIONS

MITRE performed a detailed literature search, reviewed over 100 documents, and prepared abstracts for 62 of the most relevant of these. Similarly, MITRE interviewed over 45 Automated Data Processing (ADP) and information specialists from the DOD, academia, and the corporate sector and prepared 32 interview reports of these conversations. Conclusions arising from these research efforts are as follows:

- Data redundancy is perceived as a significant problem from the perspective of both the system developer and system user. Data standardization and data dictionary efforts are underway in many military and corporate organizations to deal with the problem. This includes development and use of specialized data administration software to cope with many of the design and operational aspects of the problem. The effort has been a mixed success--ranging from significant reduction in redundant data elements to lengthy, ineffective efforts yielding minute progress.
- There is general agreement among the specialists that MITRE interviewed, that the problem of data redundancy and inconsistencies has its origin in the proliferation of uncoordinated information systems, and that an important element in problem resolution is a move to a "shared data" environment.
- There is similar, near unanimous agreement that the value of good data and information to users is the ultimate criterion for the justification of automated information systems. There is similar agreement that the effort required to quantify that value ranges from extremely difficult to impossible. Once again, everyone agrees that it is essential that metrics be developed for assessing value in use; approaches that require considerable interaction with users appear to be the most viable.
- MITRE examined efficiency-based Cost-Benefit Analysis (CBA) and effectiveness-based (Decision/Game Theory) approaches to this problem, but selected a synthesis of these approaches, called ILCA as more applicable and adaptable to Army information system evaluations. Although not a formal methodology, ILCA blends the essential elements of the other two approaches, thus providing the necessarily broader perspective required to assess the life cycle impact of information. Moreover, it directly addresses the key issue of information system evaluation--its value in use--by specifically recognizing that how users use or value information

in reaching decisions and the value of decision outcomes are uncertainties that can only be resolved through direct interaction and guided discussion with those users.

- A strawman, conceptual ILCA framework, that sketches out some of the prospective system costs and benefits that would be necessary to evaluate the impact of data redundancy in stove-pipe and shared data environments, was developed and is shown in Table ES-2.

The literature search and discussion revealed that both data redundancy and information value are areas of deep concern to the information systems community, which is reflected in the number of papers appearing in the trade journals. A general recommendation is to closely monitor the major periodicals cited in this report for new developments. For the same reason, close liaison should be maintained with many of the information specialists interviewed for this project. Specifically, contact should be maintained with the staffs of the Personnel Systems Informations Systems Command (PERSINCOM) and Naval Personnel Research and Development (R&D) Center who are embarked on reduction of data redundancy and user metrics projects, respectively.

Finally, it is recommended that the usefulness and feasibility of the ILCA approach be assessed in a specific application or demonstration project. A candidate information system could be selected from the PERSINCOM or Standard Army Management Information System (STAMIS) systems for evaluation, including a data collection effort centered on the identification and valuation of user perceptions.

TABLE ES-2
PROPOSED REPORT FORMAT OF NPV COSTS AND BENEFITS FOR
ARMY STOVE-PIPE VERSUS SHARED DATE ENVIRONMENTS

<u>Costs</u>	(Thousands of 1989 \$)	
	<u>Stove-Pipe</u>	<u>Shared</u>
Facilities	\$	\$
Hardware		
Processors		
Communications		
Peripherals		
Work Stations		
Software		
Development		
System Concept		
System/Software Requirements		
Preliminary Design		
Detail Design		
Code/Test		
System T&E		
Quality Assurance		
Documentation		
Maintenance		
Facilities		
Hardware		
Software		
Operations		
Personnel		
ADP Supplies		
System Life Cycle Costs	\$ A	\$ B
<u>User Benefits</u>		
Net Effect on Decision-making Process		
- Quality of Output Contents		
- Quality of Output Form		
- Timeliness		
- Quality of System as an Aid to Problem Solving		
- System Stability		
- System Flexibility		
Impact on Decision Outcome		
- Value of Objective 1		
- Value of Objective n		
Value of Data to Decision-makers	\$ C	\$ D
Net Present Value	\$ C-A	\$ D-B

SECTION 1

INTRODUCTION

One of the major issues facing the Army today is the need to share data in an automated fashion. In the past, Army applications were developed independently to support only the functional proponent. Little regard was given to the support that particular information systems could provide other functions. This situation resulted in redundant data files, inconsistent data, and inaccessibility of data from the many operational applications in the various Army functions. Data were typically locked into a proponent's application which has limited data compatibility and inhibited the Army's ability to share data. The inability to share data or operate in a more standardized data environment has broad implications--not only for information system efficiency but, also, for the Army's reliance on these systems in its decision-making activities.

This section of the report discusses the background of this problem, what the Army has done about it, the purpose of this research effort, and a description of MITRE's research approach.

1.1 BACKGROUND

Since 1982, the Army has been focusing increasing attention on managing information as a resource. In May 1984, the Army established an Information Mission Area (IMA) to ensure the integration of data resources. An organizational structure was also created to support this objective. The office of the Directorate of Information Systems for Command, Control, Communications, and Computers (DISC4) is the Army Headquarters component responsible for oversight and policy development for the IMA. The Information Systems Command (ISC) was created for the implementation of the IMA; and the Information System Engineering Command (ISEC), a subordinate command under ISC, was delegated the responsibility for technical development of several IMA functions. Development of an Army Data Management Program and an Army Corporate Data Base (ACDB) were two of the major functions assigned to the ISEC.

These two initiatives were designed to overcome the problem of data proliferation and confusion developed by separate Army components. Over the past several years, each of the many Army components has developed its own information systems without regard to whether the required data were available from other automated sources. This proliferation of separate information systems has resulted in many situations where Army decision-makers obtain conflicting information which makes decision-making confusing, time consuming and prone to errors. Furthermore, since much of the same

basic information is collected, maintained and processed by different organizations, the situation as a whole is inefficient and costly. The Army's basic strategy for meeting the objectives of the IMA is to formalize the process by which new systems are built, and to standardize data elements. This strategy is specified in Army Regulation (AR) 25-1, The Army Information Program, and the process to standardize data is guided by AR 25-9, The Army Data Management and Standards Program. These are the mechanisms by which the Army ensures data integration, sharing, standardization, interoperability, timeliness, and validity of information provided to Army decision-makers. These programs provide Army leadership with tools for identifying existing information resources, validating and satisfying known information requirements, and providing a systematic approach for acquiring future resources.

The ACDB concept was one alternative for meeting the objectives of the IMA. It was designed to further the aim of data sharing and to reduce redundant data caused by multiple data bases. The initiative was abandoned because of the lack of quantitative data to justify the projected expense. Many people in Army leadership intuitively "knew" that a shared data base environment was better than the many separate, disparate data base environments, but it seems that no one has specifically quantified the benefits. In particular, no one has performed cost and benefit analysis of the alternatives. This situation is the major motivation for this research effort.

1.2 PURPOSE OF RESEARCH EFFORT

The primary purpose of this report is to summarize research on the quantification of the benefits of a shared data environment (or reduced data redundancy) within either military or commercial information systems, and the effect on decision-making. This effort includes a preliminary assessment of the applicability of available evaluation methodologies to an Army data environment.

A more general definition of the research effort, often encountered in both the literature and discussions, is that the quality of data or information affects the quality of decisions. The underlying presumption (not scientifically validated but claimed by information practitioners) is that non-shared data environments are synonymous with higher levels of non-standardized, inconsistent data, and are thereby more likely to lead to faulty decision-making. Therefore, MITRE's efforts were directed at reviewing the cost and benefit issue in both guises--as a shared versus non-shared data environment problem and as a high versus low data redundancy problem. Whichever perspective is used, the evaluation methodologies were examined with respect to both Management Information System's (MISs) operational variables (data storage, program maintenance, etc.) and user/decision-maker effectiveness criteria.

1.3 RESEARCH APPROACH

MITRE first performed a detailed literature search in military, commercial, and academic areas. The primary focus of this effort was to determine if anyone had applied quantitative techniques toward measuring and comparing the effectiveness of information systems--including the benefits and costs of developing and operating a shared data environment versus a group of separate data bases. Each document related to the topic was reviewed and an abstract was written. MITRE reviewed over 100 documents and prepared abstracts for over 60. These abstracts are contained in Appendix A.

MITRE also interviewed Automatic Data Processing (ADP) and information specialists from the Department of Defense (DOD), academia, and in the commercial sector. The purpose of these interviews was to obtain background data and to collect applicable research documents. The content of these interviews are contained in Appendix B.

After reviewing and abstracting the salient information from the literature, the information was summarized and analyzed. This information is contained in Section 2. Analysis consisted of evaluating the methodology used to address the cost and value of processed data in general, and the cost and benefits of a shared data environment versus a non-shared data environment in particular. Any relevant methodologies discovered through the literature search were assessed for applicability to the Army environment. Conclusions and recommendations were then formulated, and are contained in Section 3.

SECTION 2

FINDINGS AND ANALYSIS

This section discusses the results of MITRE's literature search concerning the costs and benefits of a shared data environment over a non-shared environment. Implicit in this statement are the underlying questions of how the quality of data is affected in those two environments, and how data quality affects the quality of Army decisions. This section first presents a framework in which the findings of the literature search should be viewed. This is done by discussing the fundamental problem, the resulting questions which need answering, and by defining a few key terms. Subsequent sections summarize the findings of MITRE's literature search, and how these findings "fit" within the framework.

2.1 RESEARCH FRAMEWORK

In order to place MITRE's research findings in perspective, it is necessary to clearly state the problem and the questions MITRE is addressing, and then define certain terms.

2.1.1 Problem Statement

There are four basic components to the definition of the problem at hand:

1. The Army stores and maintains redundant data within and between Army systems because functional proponents have developed their own data bases to serve specific purposes. These are called "stove-pipe systems", and are defined in Section 2.1.3. This situation is believed to be costly and inefficient. Although the systems have been developed independently with each having its own unique hardware configurations, software programs, and data characteristics, technology can solve the hardware and software interoperability; but data compatibility problems are largely unresolved.
2. Because systems have been implemented in a non-coordinated, i.e., "stove-pipe", fashion, data elements within and between Army systems are not standard. A given data element may exist in two systems with the same name, but have different meanings. Similarly, two systems may both have the same data element, but call it by two different names. This situation is also believed to be costly and inefficient.

3. Data which are non-standard with respect to definition and composition, currency, or value, and which appear in multiple MISs can lead to poor decision-making with ripple effects throughout the decision-maker/user community. Poor decisions may lead to improper allocation of resources and concomitant cost, schedule, and technical performance penalties.
4. Army leadership believes that the problem of data redundancy and inconsistencies can be overcome by eliminating the development of systems being developed in a vacuum and moving to a "shared" data environment.

2.1.2 Research Questions

The primary thrust of this research effort is to answer the following questions:

1. Has anyone attempted to determine the costs and benefits of a "shared" data environment versus a non-shared (stove-pipe) environment in any MIS setting, either commercial or military? If so, what are the results?
2. Has anyone attempted to determine the "cost" of redundant and inconsistent data in terms of adverse impact on decision-making? If so, what are the findings?

A secondary purpose of this research effort is to determine the extent of data redundancy within MISs. It is believed that this problem has been fairly well researched and documented, particularly within the Army data environment. This report states the extent of data redundancy as reported by other researchers.

2.1.3 Definitions of Key Terms

To facilitate understanding of the findings presented in this report, it is necessary to define certain key terms and concepts. These are presented in alphabetical order in this subsection. A complete glossary of terms used in this report is contained in the glossary.

2.1.3.1 Data Inconsistency

Data inconsistency usually refers to a situation in which the same named data element (or alias) in different information systems either do not contain the same information or possess different formats.

2.1.3.2 Data Redundancy

Data redundancy is a state in which data are defined and/or maintained by more than one source. The redundant copies of the data may be in the same or different forms. Data redundancy may be planned (for performance considerations in one system) or unplanned. For the Army environment, redundant data arise because many information systems were developed for a given functional proponent without regard to possible existence of other data sources.

2.1.3.3 Integration

The office of the DISC4 defines four types of integration: functional, technical, programmatic, and program and resource. Functional integration refers to those activities that are associated with the definition and synchronization of functional requirements as a basis for accomplishing information interchange. Technical integration refers to the activities that are associated with establishing or enhancing integration or information among component parts based upon a physical architecture. Programmatic integration involves synchronizing the acquisition, fielding, and support for interdependent information systems. Program and resource integration deals with activities associated with the allocation of resources to ensure appropriate resource levels for program execution.

Frequently, integration in the Army is referred to as being either vertical or horizontal. These terms refer to function integration; the concept is depicted in Figure 2-1. Vertical integration can often be achieved through the efforts of a single functional proponent and developer team. Lowering life cycle cost or more efficient movement of information will be the primary goals of vertical information.

Horizontal, or cross-functional, integration occurs within a given Army echelon to provide echelon commands with a cross-functional view of information. The need for horizontal integration is first driven by the need for a commander to possess data across functional areas. That is, if operational effectiveness is improved, then horizontal integration should occur. However, if two information systems within different functional areas are providing the same data, then from a cost effectiveness viewpoint, functional integration should be carefully analyzed.

2.1.3.4 Shared Data Environment

The term "shared data environment" usually refers to a type of technical integration within an organizational setting. The Army has defined four levels of integration. These levels are illustrated in Figure 2-2. It is important to understand the difference between these levels

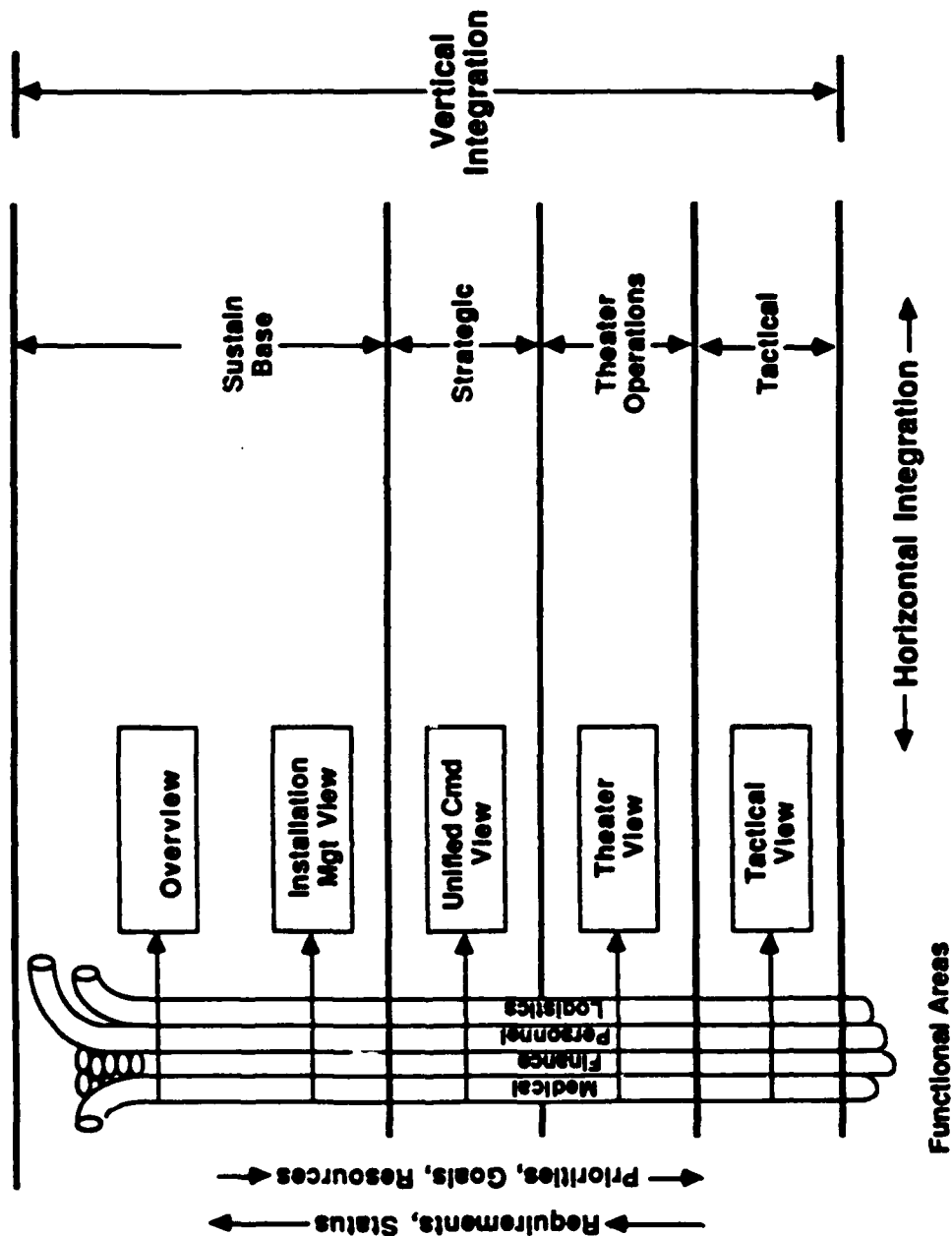
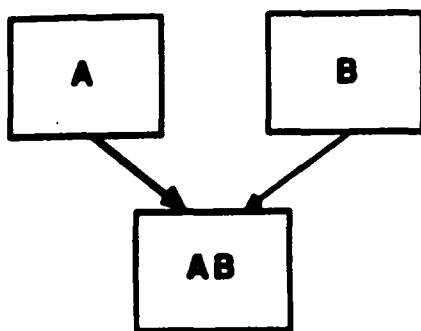
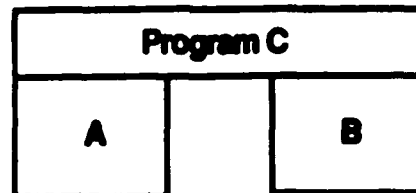


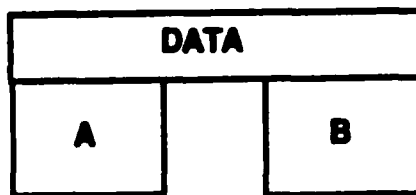
FIGURE 2-1
HORIZONTAL AND VERTICAL INTEGRATION
(Source: Mendenhall, 1988)



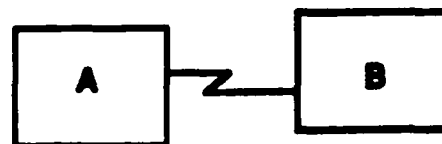
Single System



Shared System



Shared Data Base



Information Exchange

FIGURE 2-2
LEVELS OF TECHNICAL INTEGRATION
 (Source: Mendenhall, 1988)

because the complexity (and hence cost) differ drastically. The first level is a single system where the functionality and data of two separate systems are integrated. The second level occurs when two systems are designed as separate systems, but the user is able to get a unified view of the system through a common control program (shared system). The third level occurs when two systems are designed to share a common logical data base, but do not share functionality. The fourth, and crudest, level is really data interchange. Two systems (A and B) do not share a common logical data base or common functionality--data are merely exchanged.

2.1.3.5 Stove-Pipe System

A "stove-pipe" system is an information system that operates across Army environments but does not communicate with other information systems. Within an Army's functional area, information of resources and priorities, and goals for effective allocation flow down the organization; information responses such as requirements and status information flow upward. Figure 2-1 also illustrates a stove-pipe system.

2.2 THE COSTS AND BENEFITS OF DATA

This section of the report presents the findings about the costs and benefits of a shared data environment as compared to a single system (or stove-pipe) environment. Section 2.2.1 discusses the costs of data redundancy within single systems while Section 2.2.2 presents what is known about the costs and benefits of a shared data environment over a non-shared, or stove-pipe, environment. Briefly stated, very little quantitative work has been done on this topic. However, an Army application is discussed at some length. Section 2.2.3 presents a review of methodologies and techniques for determining the value of data/information in automated information systems. By focusing on the means of valuing benefits, especially as realized by users of the system, the stage is set for development of a workable framework for evaluating costs and benefits of shared and non-shared data environments. This framework is presented in Section 2.2.4.

2.2.1 The Cost of Data Redundancy in Single Systems

Understanding the problems of data redundancy and the desire to share data among users first requires an understanding of data redundancy within a single system.

As defined earlier, data redundancy is a state in which data are defined and/or maintained by more than one source. Data redundancy has come to imply a series of problems to different researchers. To some, data

redundancy is exactly that: where the same named data element is replicated elsewhere in a data base scheme. Some degree of this type of redundancy is necessary in a single, non-shared data environment. The reason for this is to give better access time and to provide the capability to recover from accidental loss of data. This type of data redundancy is sometimes referred to as "controlled redundancy" (Lobley, 1981). This type of redundancy is generally sought to be kept to a minimum because it tends to cause synchronization errors as well as "data inconsistencies". Some researchers use data redundancy to mean data inconsistencies. As defined earlier, this means the same data element being called different names or different data elements called the same name. To others, data redundancy means confusion over data element aliases or synonyms. For the purpose of this report, data redundancy means replication of a data element or some form which can cause confusion over meaning or value.

Data redundancy is often thought to only occur in several stove-pipe systems in which different people have designed and developed different data bases and called and/or formatted the same basic data element differently. However, data redundancy problems also occur in single systems (or a non-shared data environment). The nature and cost of data redundancy is discussed below.

Most organizations have a great deal of data redundancy: the degree of redundancy has been documented in many organizations as close to 75 percent. [See the abstract "Putting Top-Down and Bottom-Up Analysis Together" (Rice & Laufer) in Appendix A, and discussions with LTC Hollist of the Personnel Systems Information Systems Command (PERSINCOM), and Mr. Nguyen from the Air Force Staff/Systems Command, Control, Communications and Computers (C4) contained in Appendix B.] Redundant data in a single systems have the following cost impact:

- Take up more disk space than required and therefore more disk space is purchased or leased
- Magnetic media or removable disks cost is higher because more backup media is needed
- Access time is reduced because the disk heads must read more data
- Program development time is longer (and hence cost is greater) (Lobley, 1981)
- Program maintenance costs are higher because of data naming issues (Rice & Laufer, 1988)

No research products were found which rigorously measured the extent of data redundancy cost, but it is not difficult to quantify. One set of researchers estimates that most of the cost of data problems is in software

maintenance. They estimate that 50 to 80 percent of data processing departments personnel budget goes toward software maintenance and that probably 50 percent of the maintenance effort is attributable to bad coding or bad data naming (Rice & Laufer, 1988). Briefly stated, the biggest reason cited in the literature for bad data is a lack of planning and control over data definitions. Two major concepts, if properly applied, can help exercise some form of configuration management and mitigate data redundancy. The first is the idea of a data administrator (DA). This is a person (or organization) responsible for data planning and control. The second concept is that of an automated data dictionary.

The DA reviews and approves all data base design, maintains and assigns code values; assigns and controls security of data files; and establishes and enforces standards for documentation, data base design and system controls.

A data dictionary is a repository for definitions and related information for data resources of an organization. It is typically a software tool used for the development and maintenance of application systems. It maintains and controls the use of all data definitions within application systems. It contains descriptive information, technical characteristics and the interrelationships between data elements. Using an automated data element dictionary will reduce data redundancy and consequently the cost and time to develop and maintain an information system (Lobley, 1981).

A strong DA, coupled with a good data dictionary is the only way to reduce data redundancy and standardize data element meanings to all concerned. However, to be effective, the whole software creation process, from initial planning through design and code generation to maintenance, must be automated. Computer Aided Software Engineering (CASE) tools are developing to the point where the tools become part of the data administrator "policing" mechanism. CASE automatically enforces the procedures surrounding a data base each time a programmer works on a new application or maintains an old one. Benefits of the overall data administration have been estimated by several researchers who project savings of 10 to 15 percent in development and 80 percent in systems maintenance (Voell, 1986).

2.2.2 The Cost and Benefits of a Shared Data Environment

This section further defines the meaning of a shared data environment, the benefits of a shared data environment and briefly describes the cost and benefit results achieved in industry's and the Army's efforts of moving to a shared data environment.

2.2.2.1 Data Sharing

As stated earlier, a shared data environment has different meanings. It can range from merely electronically passing data from one system to another to sharing a common logical data base. The term data sharing also connotes Distributed Data Processing (DDP). Simply stated, DDP is a system involving multiple sites connected in a communications network, in which a user at any site can access data at any other site. However, the term is ambiguous. It has been used to refer to highly centralized computing systems with remote terminals at one extreme, to totally decentralized connections of stand alone minicomputer systems on the other. It may include distribution of processing to remote sites, distribution of data collection, distribution of applications development, distribution of software programming, distribution of data itself, distribution of computer operations, and distribution of the control of computing resource. Each of these can be more or less centralized or decentralized depending upon organizational requirements.

For purposes of this report, when data sharing is discussed, we are referring to some form of Information Sharing (IS) which involves the sharing of resources (data, equipment, or personnel) between two or more organizations. The key point to note is that the cost of IS is driven by which resources are shared and how the sharing is implemented.

2.2.2.2 Potential Benefits

The primary reason for an organization to move to an IS environment is to reduce cost and to improve performance.

Within the commercial sector, the following advantages for moving to an IS environment are believed to be valid:

- Return on investment will be increased due to lower costs
- Better information will permit the firm to achieve a more competitive edge, maintain market leadership or enter a new product/market area
- Software development cost will be shared
- Maintenance costs will be shared
- System reliability will be increased
- Operating costs will be shared
- Resource allocation decisions will be more efficient

The benefits of an IS environment within the Army have been discussed over the last several years. Some of the important beliefs and hypotheses are as follows:

- Many functional components need and use data which are also generated by other functional components. Therefore, sharing of common data will be more efficient and hence save money by eliminating duplicate resources (e.g., hardware, software, personnel, data collection and maintenance efforts).
- Data sharing will result in consistent and accurate data, and thus will significantly improve the allocation of resources and hence readiness.
- Data sharing will result in timely, synchronized data being used in the decision-making process and this will improve the quality of decisions made.
- Data sharing will result in data standardization which will markedly improve the ability of Army personnel to perform their jobs (Pueschel, 1988).

2.2.2.3 Valuing the Information Sharing Experience: Some Applications

In the commercial sector, one researcher found that when a wholesale distributor moved to an IS environment, the firm's financial picture improved to the extent of providing a 30 percent return on investment. The major changes which affected the turnaround were an average 40 percent reduction in order/buying personnel, an increase in market share from 22 to 30 percent and a reduction in Electronic Data Processing (EDP) costs. Furthermore, the firm noted a decrease in error rates which led to administrative cost avoidance or reduction (the processing of fewer exceptions per message) (Barrett, 1982).

In spite of such gains, many IS efforts fail. Capers Jones, founder of Software Productivity Research, claims that 25 percent of all large real-time and telecommunications-based software projects are never completed. He notes that the big projects fail more frequently than the smaller ones. The Massachusetts Institute of Technology conducted a data management practices study of 20 major companies and found that five companies attempted an IS environment and all failed. The reasons cited for those failures centered on problems with the meaning of data. An example is the distinction between a sales order and a purchase order. Should the same data code be used to represent a customer who is also a supplier? If the same code is applied, how would the difference between the customer-vendor and the vendor-customer relationship be represented? (Van Rensselaer, 1988)

Several other reasons for these IS development failures are offered, such as volatile business conditions, poor quarterly performance measures, competition, primitive hardware access methods, and sloppy programming. However, several researches believe that arcane data structures, and data inconsistencies are the major stumbling block (Carlyle, 1983 and Van Rensselaer, 1988). Within the Army, very little research comparing the cost and benefits of an IS environment to the single systems (or stove-pipe systems) appears to have been done. However, one cost benefit study of the various alternatives for technical integration of the Army's personnel components was conducted. The results of this study are briefly discussed below.

In an effort to promote data commonality among the Army Military Personnel Center (MILPERCEN), the Civilian Personnel Center (CIVPERCEN), the Army Reserve Personnel Center (ARPERCEN), and the National Guard Bureau (NGB), the Army initiated development of a Total Army Personnel Data Base (TAPDB) that will include a networking capability linking the four components. TAPDB, when fully implemented, will provide standard data elements and a common data structure to enable exchange of personnel data among the components over a communications network. The various components, at the time of the initial study (1986), exchanged data only minimally. This exchange occurred primarily through tape transfers and was inefficient and time consuming.

Three basic alternatives were analyzed and compared to the baseline alternative. The alternatives are as follows:

- Baseline--Transfer of tapes between components
- Remote Access
- Transparent Heterogeneous Access
- Transparent Homogeneous Access

The remote access alternative involves connecting all four components via the Defense Data Network (DDN). Each user on the network will use communication software to access another component's host computer. The host computer treats the remote user just as any other terminal on the system, i.e., the remote user must follow the conventions and procedures established by the host software.

The transparent access alternatives (heterogeneous and homogeneous), are identical to the remote access except that a single set of conventions and procedures will be used to access data at all components; and, the data base management system will automatically transfer data between components on demand so that users will not have to be concerned with the conventions and procedures of the remote site or the physical location of the data being

accessed. The heterogeneous alternative involves different hardware and software. Specific software must be written to provide session control, merge/extract functions, perform data representation conversions, and provide application linkage to the communications subsystem. The homogeneous alternative does not require software development beyond the linkage to the communications subsystem. It is a special case of the heterogeneous transparent access where the computer and the Data Base Management System (DBMS) are the same. In the homogeneous method, the merge/extract and convert functions become unnecessary, and only a simple link to the communications subsystem is necessary.

The results of this cost benefit analysis were presented in terms of cost/benefit (C/B) ratios and payback periods; these data are summarized in Table 2-1 below.

TABLE 2-1
RESULTS OF TAPDB COST BENEFIT ANALYSIS
INTER-COMPONENT DATA ACCESS ALTERNATIVES

<u>ALTERNATIVE</u>	<u>C/B RATIO</u>	<u>PAYBACK PERIOD</u>
Remote Access	3.31	3 Years
Transparent--Heterogeneous	1.26	6 Years
Transparent--Homogeneous	2.24	4 Years

The analytic technique used in this effort was the incremental improvement method. This technique consists of calculating the costs associated with obtaining an incremental improvement. Benefits were defined as cost reductions, future cost avoidance, or improved quality or availability of a new capability. Each alternative's costs and benefits were compared with the baseline, and only the respective differences are presented. Incremental costs included hardware, software, maintenance, start-up, operations, and training.

The benefits were identified as tangible or intangible. Tangible benefits were those that represented a quantifiable cost savings or avoidance to the implementing organizations as a direct result of developing and utilizing the improved method. Intangible benefits are those benefits which result in genuine improvements in the system but cannot be quantified. The tangible benefits identified were:

- Reduction of personnel and ADP expenditures to transmit and process other components' data
- Reduced personnel resources expended to fully audit transferred records

- Improved retention rates

Some of the intangible benefits identified (but not quantified) were better access to data, improved response time, facilitation of data standardization, and improved retention rates. Improved response time was believed to be particularly important because it would permit more accurate strength assessments, more accurate and timely mobilization processing and more timely notification of potential recruiting candidates. Additionally, timely sharing of potential personnel losses and the sharing of knowledge of vacancies before separation occurred would improve the retention rates within the Army and prevent the loss of substantial investment in trained soldiers.

The results of the TAPDB are important, but the study did not go far enough into the benefits of a shared environment. Many of the primary reasons for sharing data were categorized into the "intangible" category, and were not quantified. Because it was beyond the scope of the effort, the study did not address the cost and benefits of no data sharing. More can and should be done to quantify the benefits of having better data which can accrue in an IS environment.

After reviewing and assessing the general approaches for quantifying system and decision-making benefits in the next section, a conceptual framework for cost and benefit evaluation of shared versus non-shared data environments will be presented.

2.2.3 The Value of Data and Information

This section examines the potential for assessing data quality with respect to its effect on decision-making. As such, it carries the previous analysis of data quality and redundancy in stove-pipe and shared data environments to the next logical linkage, decision-maker/user performance.

2.2.3.1 Background

This background review provides the contextual setting and alternate definitions of the quality assessment problem; its perceived significance in the information systems community; the heightened focus on information resources; and the major research direction for coping with quality assessment and measurement.

2.2.3.1.1 Problem Definition. The specific problem under investigation in this task can be stated in a number of analogous ways:

- To what measurable extent does bad data lead to poor decisions?

- What are the value of standardized, non-redundant data to automated information system managers and users?
- What are the cost and benefit criteria used to justify design/development/operation of automated information systems for decision-making?
- What is the value of information? The essence of the data economics task, however expressed, is encompassed in the following considerations:
- Evaluation of information systems needs to include both system efficiency (process performance) and user effectiveness (user performance) aspects
- Evaluation of user or decision-maker effectiveness requires identification of benefits accruing to improved data or information, and techniques for measuring or assessing those benefits

2.2.3.1.2 Decision-Making Environment. As a companion piece to the overview of the data environment presented earlier, it is useful to provide a brief sketch of the decision-making environment. Decision-making has been defined as "the conversion of information into action" (Taylor, 1986). Information, in the context of this report, is used interchangeably with data, although strictly speaking, information is processed data (i.e., it is the output of the automated information system, where data are input). However, data and information can both be looked as input to the decision-making process.

Whereas the quality of data is susceptible to measurement (e.g., accuracy and recency), the quality of decisions is highly subjective. Even subjectively, however, a decision can be evaluated and often is. Moreover, because of the frequent time lag between the decision point and the consequences of that decision, ensuing events not known or predictable by the decision-maker may have a greater impact on the outcome than the actual decision. Because causality is blurred, therefore, good decisions are often defined as those that follow the proper decision-making procedures. The contention is that following the procedures is more likely to result in the attainment of the decision-maker's objectives in the long run and have the best consequences because he has engaged in diligent information processing. These procedures (Taylor, 1986) include the following:

- Proper definition of the problem
- Identification of wide range of alternatives

- Survey of the full range of objectives and the values implied by the choice
- Careful consideration of costs/risks of consequences of each alternative
- Search for new information to expand evaluation
- Assimilation of new information/expert judgments, even when not supportive of preferred action
- Re-examination of positive/negative consequences in light of all information
- Provision for implementation plan and contingency plans if risks materialize

Getting a passing grade on these criteria ensures a higher probability of successful decision-making. This does not mean that good decision-making can be judged solely on the basis of ex ante behavior, with no need to review consequences. In fact, several of the criteria focus on costs, risks, and value of goal fulfillment--which, collectively, are the basis for an ex post cost/value assessment of decision quality. Nonetheless, our review of the literature shows that--while there is a near unanimous consensus on the subjectivity of decision-making--methodologies exist for converting subjective measures of goal fulfillment, revealed preferences/utilities, user perceptions, and risks into quantitative measures. These are discussed below.

Finally, there are two additional features of decision-making that affect the assessment/measurement of outcomes:

- **Multi-attribute Decisions/Multiple Decision-Makers.** Very few decision outcomes can be characterized by a single number. It is more likely that there will be a number of properties, attributes, or objectives met by which the decision is to be assessed. Moreover, they are likely to be of varying importance to the decision-maker as well as expressed in different units (e.g., time or dollars). Assuming that the awesome task of obtaining some measurements for all these attributes--and for all decision-makers/users--can be performed, the complications of weighting and incommensurability add such a burden in time and energy that most decision-makers do not optimize--they "satisfice", i.e., look for solutions that are acceptable, reasonable, or satisfactory after a moderate review of the options (Hogarth, 1987).
- **Programmable versus Non-Programmable Decisions (Taylor, 1986).** Programmed decisions are well structured, routine, repetitive, and can be generally based on rules, standard operating procedures,

and simple computation methods. The amount and scope of information can, in general, be anticipated. Information needs are routine and well defined, for the most part quantitative or factual. There are clear cut information channels and structures. Management financial, and production information systems are generally considered programmable. These types of systems often are used to monitor and detect current differences with a pre-set level of desired performance. Data, accuracy, currency, reliability and validity are highly valued in this environment.

On the other hand, non-programmable decisions tend to reflect quite different criteria for judging the value of information. This is due to the uniqueness of the problems faced, the lack of predictable structure, greater uncertainty, and the heuristic techniques needed for solving the decision problem. Much of the information is informal and verbal. Selectivity of data to delimit information overload is critical here; therefore, it is of higher value than the same trait in programmable decisions.

The different emphasis placed on the value of specific data quality attributes in programmable and non-programmable decisions is not meant to be a definitive distinction, but suggestive of likely relative evaluations of decision-makers in the two scenarios.

2.2.3.1.3 Significance of the Problem. It has been repeatedly and emphatically claimed both in the literature and in discussion with information system specialists that the problem of determining the value of information systems is the central issue facing the information technology world, especially since the relatively unchecked information explosion which has hitherto taken place will in the future be subjected to more careful economic scrutiny and budget limitations (U.S. Congress, 1980; Office of Management and Budget (OMB), 1986, 1988). To the extent that this expansion in systems has been justified, it has been based on cost savings over the current operation. Any zero-based evaluation would pose the more fundamental question: why is any information system necessary? That translates into the central theme of the data economics task, i.e., measuring or assessing value or benefits accruing to users or showing how information adds value to the performance or effectiveness of the user (decision-maker) community. Even in the absence of a zero-based requirement, any system that cannot adequately justify its development on the basis of internal system efficiency, is well advised to carefully review, assess and incorporate potential user benefits in its evaluation scheme.

2.2.3.1.4 Role of Information Resources Management. Recognition of the fact that information must be managed as a valuable and costly resource has been slow and piecemeal. The paperwork Reduction Act of 1980 (U.S. Congress, 1980) formally set the framework for what is called Information

Resources Management (IRM) in the federal government. IRM is "an approach to applying appropriate and effective management philosophy, methodology, and techniques to decisions about data and information and other information resources (equipment, software, personnel, etc.)" (Chick, 1986). The underlying objectives of IRM are to assure that information produced from information resources has maximum value and is produced at the most efficient cost. The emphasis on value has spawned a large number of papers dealing with user perspectives by academics and information system specialists--most of which point to the IRM concept as the triggering mechanism.

Although there is no IRM "value methodology", IRM advocates a method called information environment analysis, i.e., examination of the environment in terms of such variables as organization objectives, management communications, and types of decision-making processes (Levitan, 1982). The OMB, in its information technology oversight role prescribed by The Paperwork Reduction Act, complements IRM's information analysis emphasis by requiring cost-benefit analysis for all information technology initiatives whose total cost exceeds five million (M) dollars (or \$2M in one year) or which represents--in OMB's view--a significant information system. Specific guidelines for the analysis include demonstrating a 10 percent return on investment or--failing that--providing additional justification in terms of quantifiable but non-economic improvement to the requesting agency's ability to perform its mission (OMB, 1988). It seems clear that information value in decision-making is an issue receiving increasing attention at the federal level.

2.2.3.1.5 Survey of Methodological Approaches. The range of methodologies for assessing the value of information in information systems reflects the shifts in the perceived mission or challenge for the information community. These can be classified in terms of the three information issues (Black, 1982):

- Quality versus quantity of data
- Functional versus process orientation
- Transfer versus content of information

Differences in the emphasis placed on these issues by information specialists helps explain the resultant differences in methodological approaches. For example, the increasing demand for information over the years has fostered technological innovations to continually expand the capacity necessary to store, retrieve, and process data. There is a large segment of the information community that accepts this continued growth of information, and seeks to meet (and justify) the demand for greater information technology to accommodate the data. At the other end of the

spectrum are those who argue that meeting the challenge for an increase in the quantity of information is not the solution, but may in fact compound the problem. This view stresses the importance of the quality of information--a limited set of relevant and accurate data--which could go further to solve problems than would a glut of information. Moreover, this data quality group is focused on the content and functional use of information whereas the data quantity group is more process-oriented. This means that information value is measured in terms of the efficiency with which data processors can store, retrieve, and transfer information. This contrasts with decision-makers' perspectives which are more likely to measure value in terms of the actual use of the information, i.e., user performance or effectiveness.

Mirroring these differences in philosophy and orientation, there are two basic schools of thought on justifying investment in information technology on systems--one keying on the efficiency of the information process, i.e., the degree of cost-efficiency in combining resources to produce an information output, the other school of thought primarily concerned with the effectiveness with which the output is used to enhance organizational/user performance. The resultant approaches arising from the efficiency and effectiveness orientations are cost-benefit analysis and decision theory, respectively. The approaches are narrowly defined to imply that they look exclusively at values either in information production or information use. The narrow focus corresponds closely to the manner in which these approaches are actually used in information system analysis. A third generic approach which is referred to as information life cycle analysis combines elements of both these basic approaches and examines the costs and benefits over the entire information cycle (from production to use). The three approaches are summarized in Table 2-2. More detailed explanations, along with descriptions of selected methodological applications, are presented in the following section.

2.2.3.2 Assessment of Approaches

Brief descriptions of the cost-benefit and decision theory methodologies, along with a more detailed explanation of the hybrid approach used in the information life cycle analysis, are presented in this section. In addition, six methodological applications representing a wide range of decision-making problems found in the literature search are included to show the diversity in approaches and outcome valuations. Finally, a prospective conceptual framework for evaluating ISC information systems is presented as a preliminary check-off list of relevant cost and benefit elements.

2.2.3.2.1 Cost-Benefit Analysis. Cost-Benefit Analysis (CBA) is frequently used to evaluate information system development or modification by the criterion of cost displacement, or savings. Impact on the perfor-

mance of the organization is usually restricted to the information system environment. The costs are the costs of putting the system in place (or modifying it); the benefits are the offsetting cost reduction in information storage, processing, transfer, and management. Both costs and benefits are expressed in dollars, thereby permitting return on investment calculations.

Of course, CBA users are clearly aware of external and non-monetary, intangible benefits. Unfortunately, the difficulty of identifying and assigning values to these benefits generally leads to their omission in the analysis. When the non-quantified benefits are noted, they are usually restricted to the data processing environment--e.g.:

- Reliability and availability of the system
- Simplicity and flexibility for user needs
- Invisibility of underlying characteristics to users
- User-friendliness
- Adaptability to new requirements

As noted in Table 2-2, these "benefits" relate to the efficiency of the information storage and retrieval phases of the information cycle--not the user performance aspect.

2.2.3.2.2 Decision Theory. Decision theory deals with problems of choice or decision-making under uncertainty, i.e., where probabilities concerning alternative outcomes are unknown. It differs from game theory in two major ways. Game theory deals with choice problems under risk, i.e., where probabilities of the alternative outcomes are known or can be at least estimated. Moreover, in game theory there is a major element of predictability in the behavior of the other players since they can be expected to try to reduce the first player's payoff. In decision theory, there is no such opponent. This second player is referred to as "nature" and the corresponding decision problem is called games against nature--a nature which cannot be counted upon to act in any predictable way. The task, therefore, is always to reduce the decision theory problem to a game against a nature which does exhibit some associated--albeit subjective--probability distribution of outcomes. This is done by seeking additional information of a particular quality and quantity that can reduce the uncertainty concerning the probability of outcomes. That is why decision theory is often viewed as a statistical or probabilistic theory of the amount of information needed to reduce uncertainty in decision-making. In this context, information which helps to identify or modify these probabilities and induce a greater amount of certainty is valuable in enabling one to choose the best alternative with the best possible outcome (i.e., maximize

the decision-makers' net expected value or pay-off). Figure 2-3 is an illustration of this link between better information and payoff maximization. Information is not a free good; moreover, at some point, its cost increases at a faster rate than its value. This example of the law of diminishing returns to the increase of information results in selection of an optimal information level that maximizes net benefits or payoffs (value minus cost) that is considerably short of reducing all uncertainty or maximizing total payoff (Balakrishnan, 1988). Several interesting examples of the notion of value of information and its use in building probability distributions of possible outcomes (and thereafter assessing net benefits) are contained in RAND reports (Ginsberg, 1971 and Nelson, 1960), which deal

TABLE 2-2
OVERVIEW OF SELECTED APPROACHES FOR EVALUATION
OF INFORMATION SYSTEMS

APPROACH	ASSUMPTIONS	OBJECTIVES	INFERENCES	FUNCTION VS PROCESS	STAGE OF LIFE CYCLE	INFORMATION TRANSFER VS USE	KEY INFORMATION VALUE CRITERIA
COST-BENEFIT ANALYSIS	Increased information needs are critical. Technological innova- tions can be used to address the challenge.	The development of efficient storage and retrieval systems.	Given efficient data management, use will automatically follow.	Process	Storage and Retrieval	Transfer	Technological factors-- reliability, simplicity, flexibility, adaptability.
INFORMATION LIFE CYCLE ANALYSIS	Concern with increasing data related costs. Quality, not quantity, is the critical issue.	Improve communi- cations between users and proces- sors. Improve decision-making frameworks.	Minimize redundancy and improve quality of information. More accurate and relevant data for more prescrip- tive decision-making.	Both	All	Both	Credibility of data-- accuracy, timeliness, and relevance. Value of assessments/ judgments made by users.
DECISION THEORY	Quality of information is important. Quantity and quality are both controlled--unavailable or too costly.	The optimal amount of information to make a decision.	Given a certain quan- tity of information, a decision maker can reduce uncertainty and maximize his utility.	Function	Analysis and Use of information	Use	Relevance of data-- effectiveness in reducing uncertainty/ increasing utility. Timeliness, accuracy.

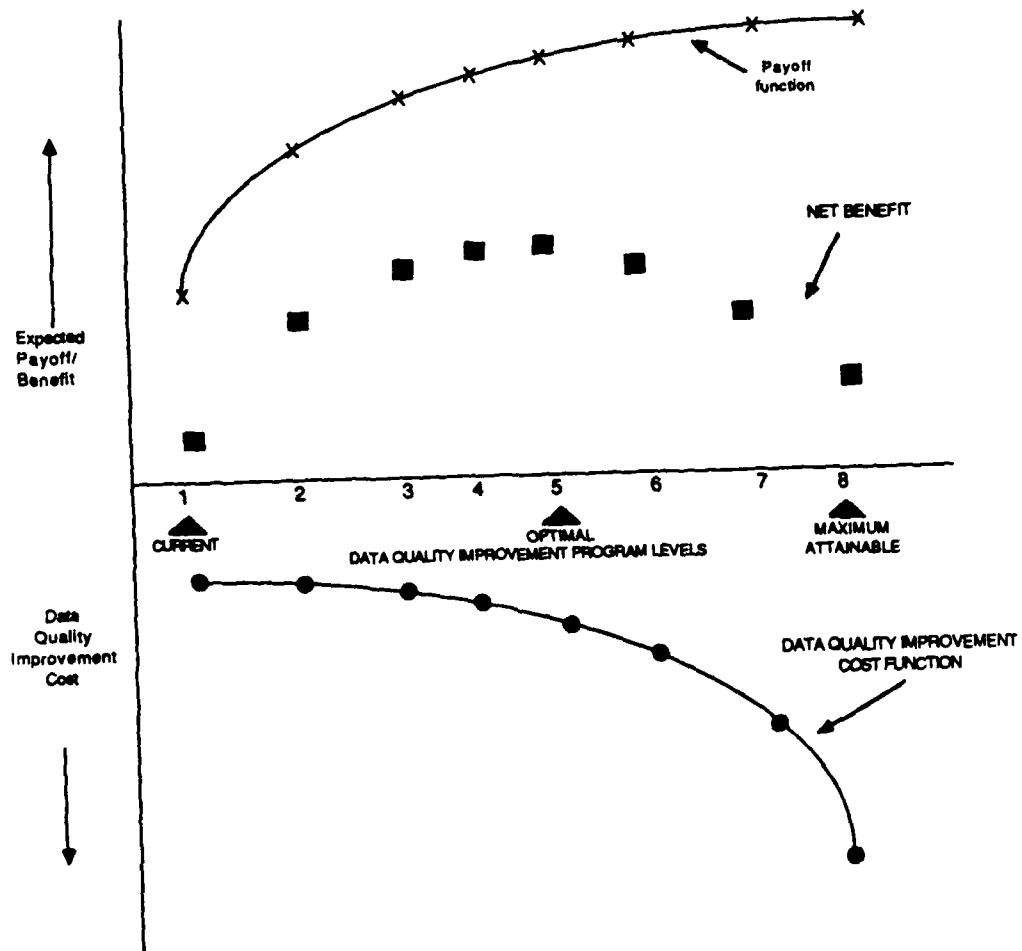


FIGURE 2-3
 PAYOFF VERSUS DATA QUALITY IMPROVEMENT COST
 (Source: Balakrishnan, 1988)

with decision modelling of medical diagnosis and weather prediction, respectively.

These payoffs, unfortunately, are based upon the assumption that the decision-maker at least knows how to measure the consequences of a decision or alternative outcome. Having solved the uncertainty phase of the problem (i.e., attributing probabilities to outcomes), one is confronted with the additional problem that outcomes generally are multi-dimensional (time savings, cost savings, etc.) and incommensurable, thereby making it impossible to express outcomes as a single number. This demands that additional value tradeoffs among the various attributes of the decision be made so that the maximization of value or utilities process reflects the decision-maker's preferences concerning those attributes (Keeney, 1976). The twin problems of reducing uncertainty about outcomes and the measurement of multi-dimension outcomes are certainly not unique to the decision theory approach. In fact, decision theory provides the analytical framework and concepts to address both. Moreover, a good case could be made for its application in the data economics task:

- Focus is on outcomes or consequences, not efficiency
- Outcomes are known (can be described in some multi-dimensional "good to bad decision" continuum)
- Probability of these outcomes is not known, but buying more information in terms of reducing data redundancy and constructing some relationships between level of redundancy and outcomes can provide some subjective probabilities

However, its complexity makes it more appropriate for sophisticated operations researchers and modelers rather than general use as an information systems application tool. More importantly, however, the methodology focuses almost exclusively on outcomes and does not explicitly bring out the internal efficiency effects of the information system itself. Since efficiency and effectiveness criteria are applied to information system evaluations, it is appropriate to develop an approach that combines the cost-benefit emphasis on efficiency and the decision theory's outcome emphasis. This third approach is discussed in the next section.

2.2.3.2.3 Information Life Cycle Analysis. This approach is a composite of the two approaches discussed previously, which focused, respectively, on the storage/retrieval and use phases of information.

2.2.3.2.3.1 Comparison With Previous Approaches. Information Life Cycle Analysis (ILCA) explicitly recognizes the interrelationship of the various stages of the information cycle: determination of information

needs, the search for and collection of data, storage and retrieval, processing, analysis, use, and disposition. Although each stage is often analyzed in isolation, decisions made at any point can severely restrict or expand the options available at each successive step. Therefore, it is preferable to have an approach that views information in the broader context of the total life cycle. For example, ILCA bolsters the cost-benefit approach in a number of important ways:

- Information system improvement that cannot be justified solely on the grounds of efficiency, i.e., system cost savings, can incorporate measures of improved managerial performance. This kind of expanded scope is part of OMB's guidance provided in Circular No. A-11 (OMB, 1988).
- An underlying assumption of the cost-benefit approach is that, given efficient storage of large amounts of data with easy access, beneficial results for decision-making will automatically follow. In fact, much of the existing information research implies that efficiency and effectiveness are indeed positively related. However a recent study, undertaken to statistically compute the association between various system efficiency measures and user-perceived effectiveness measures, found many instances of negative correlation. This led the researchers to conclude that the use of either efficiency or effectiveness measures, as a surrogate for the other, is not justified, and that evaluation of information systems requires explicit examination of both efficiency and effectiveness measures (Srinivasan, 1985).

In a similar vein, ILCA is preferred to decision theory and its concentrated focus on outcomes, and not solely because of the expanded perspective it offers. Other advantages include:

- The quality of information is more critical to ILCA than quantity --not primarily because of the unnecessary cost for excessive information or its diminishing marginal utility as viewed in decision theory--but because of the human limitations in decision-making. While information is critical, only a certain amount can be assimilated within the organization. A major reason is the limited capacity of human decision-makers to absorb and use all of the available information (i.e., information overload). Given the reliance on heuristics in human decision-making, a narrow set of salient and relevant data is often the key to better and faster decision-making rather than an overabundance of unreliable, undifferentiated data. Selected, credible information is more likely to ensure that decisions are made with more confidence and with a greater probability of anticipated results.

- The ultimate value of information depends upon the perceptual screen through which data are filtered and viewed. No matter how accurate and relevant the information, the decision-maker's perceptions, bias, and experiences--the "human baggage"--are a critical part of the decision process. Consideration of user perception and concepts of value is therefore essential in assessing outcomes. ILCA, by recommending a survey of user attitudes, recognizes the uniqueness of each decision-making scenario.

In summary, it should be noted that ILCA is not a formal methodology, but simply an approach that (1) expands the focus of the CBA and Decision Theory methodologies to include both system development perspectives of the CBA and the outcome orientation of Decision Theory (i.e., it provides a life cycle perspective), and (2) recognizes that the valuing of outcomes in information system analysis requires direct interaction with users.

2.2.3.2.3.2 Value of Information. It is generally recognized that information is not a free good. There is no such agreement, however, to its actual value. What seems clear is that information value cannot be measured in precise terms prior to use of the information. An information message only has potential for value, which ultimately is bestowed by the user. This means that the better we understand the environment in which the output of a system will be used, the better we will be able to estimate the value of those information outputs. Focusing on output of a system allows us to look backward at the system and its efficiencies, and forward to the users' choices and performance. Characterizing the problem in this manner leads to a potential formulation of the value of information, which also provides a useful perspective for the data economics task. Total value can be expressed in terms of the following function:

$$\text{Total Value} = f(A, B, C) + X$$

where

- A - Value derived from input efficiencies
- B - Value derived from how the use of information affects the decision-making process
- C - Value determined by how the decision affects the organizations's goal fulfillment (i.e. decision consequences)
- X - Unspecified aspects of value

Putting this general formulation into the context of the data economics task, the value of better information (in the form of reduced data

redundancy) is equal to the sum of its effects upon the following:

- The efficiency of the information system (e.g., cost savings)
- The efficiency and effectiveness of the decision making process (e.g., time savings)
- The decision consequences (e.g. "better" decisions)
- Other, undefined elements (optional)

The additive nature of the terms in the formula presumes that a common unit of measure, preferably dollars, will be developed. Assigning absolute, consistent, and uniform dollars to represent the value of information being produced and used by an organization will be very difficult. On the practical side, however, such denomination of information value is essential for many management purposes, including:

- Periodically confirming the continued need for information currently being produced
- Establishing priorities and allocating resources for providing new information
- Identifying problems that result in information value losses or reductions (e.g., data redundancy)
- Establishing priorities, allocating resources and establishing targets for correcting information problems that reduce or eliminate information value

Efficiency criteria relating to value (e.g., cost and time savings) are more feasible candidates for monetization than the effectiveness criteria such as decision-maker performance and decision outcomes (except for commercial decisions where profit, revenues, return on investment are quantifiable). Developing insights into those processes and outcomes requires a survey and subjective evaluation of user perceptions and assessment of impact on well-defined organization goals. Several techniques, highlighted below, are suggestive of the kind of user interaction, survey techniques, and outcome evaluation that is necessary.

2.2.3.2.3.3 Valuing User Perceptions: Some Applications. There were a number of interesting applications uncovered in the literature dealing with the valuation issue; these are identified and briefly characterized here. More detailed descriptions appear in Appendix A. These applications include:

- Matlin, 1979. Matlin assigned dollar measures of value to information systems, based on how closely the system met a derived, weighted set of corporate objectives and values of goal achievement extracted from extended discussions/interactions with corporate managers.
- Keen, 1981. Keen suggests using a technique he calls value analysis to justify decision support systems rather than trying to rigorously calculate their costs and benefits. He suggests first developing a prototype system which can be considered Research and Development (R&D) and thus does not have to be rigorously justified. After the potential benefits of the system are clearer, an assessment of the final implementation can be made, and a rigorous cost-benefit analysis done only if the estimated benefit is not obviously greater than the cost.
- Hamilton and Chervany, 1981. The authors outline a means for evaluating system effectiveness including derivation of system objectives, identification of intangible qualitative effects, accounting for dynamic changes, and differentiation of various user viewpoints. They cite several approaches to ascertain system performance including, jointly agreed system developer and system user specifications, user attitude surveys, post installation reviews, cost-benefit analysis, and Delphi procedures.
- Buck and Horton, 1988. The authors identify various elements in information systems--accuracy/currency of data; ease of information use/format; impact on organizational productivity; impact on organizational effectiveness; and impact on financial position. They evaluate a number of information systems options by rank ordering these performance elements and costs and develop cost/value indices for making "cost-effective" choices.
- Horton and Pruden, 1988. This study focuses on a benefit-cost evaluation of Department of State (DOS) information systems. Improved productivity, decision-making, performance of tasks, learning and other factors--including cost savings--were evaluated in dollars by use of a Delphi technique and validated by a sampling of senior DOS executives. A rate of return was then computed following OMB's Circular A-11.
- Srinivasan, 1985. In testing the hypothesis that user perceived effectiveness measures can and do move in the opposite direction of system efficiency criteria, the author did an extensive survey of 29 firms to measure user perceived measures of information systems such as output or report content, report form, problem solving capability, input procedures, and system stability. Perceptions were measured on a five-point scale.

- King and Rodriquez, 1978. The authors outline an evaluation process for value assessments including attitudes, value perceptions, information usage, and decision performance which can be measured at various stages of the information life cycle. A demonstration project was run to test "pre versus post" comparisons of these measures to see how perceptions vary over the life cycle.

2.2.4 A Conceptual Framework for Information System Cost and Benefit Evaluation

Based on the preceding assessment of the stove-pipe and shared data environments, the associated system costs, and the available approaches for valuing life cycle benefits, a simplified conceptual framework for evaluating alternative information system environments has been developed. The key question which must be answered is: what are the costs and the benefits of a data sharing environment and how do these costs and benefits compare to an environment where data is not shared for the same area of data interest? The issue is which alternative provides the most cost-effective solution to the Army data quality problem.

2.2.4.1 The Required Steps

In order to answer the key question above (or any cost-benefit question), requires certain analytic steps. The analyst must define the problem, perform design analysis, collect relevant data, analyze data and prepare the results (Pueschel, 1988). Only some of the required steps will be discussed here.

2.2.4.2 The Alternatives Defined

The problem statement and objectives have already been discussed. The next step is to clearly define the alternatives. The two alternative "information delivery methods" are the stove-pipe systems and some form of IS within the same functional areas. The stove-pipe system, as defined earlier, collects, processes, and passes data within an organization in a vertical fashion, i.e., from Headquarters, Department of the Army (HQDA) to the Major Commands, and to the installations and return. In an IS environment, data is shared across functional areas. Figure 2-4 illustrates how the two alternatives can deal with overlapping data areas of interest, in a hypothetical case. In the case of the stove-pipe environment, each of the three systems operates independently. In the IS alternative, the figure illustrates where data of interest might overlap and where data sharing can be beneficial.

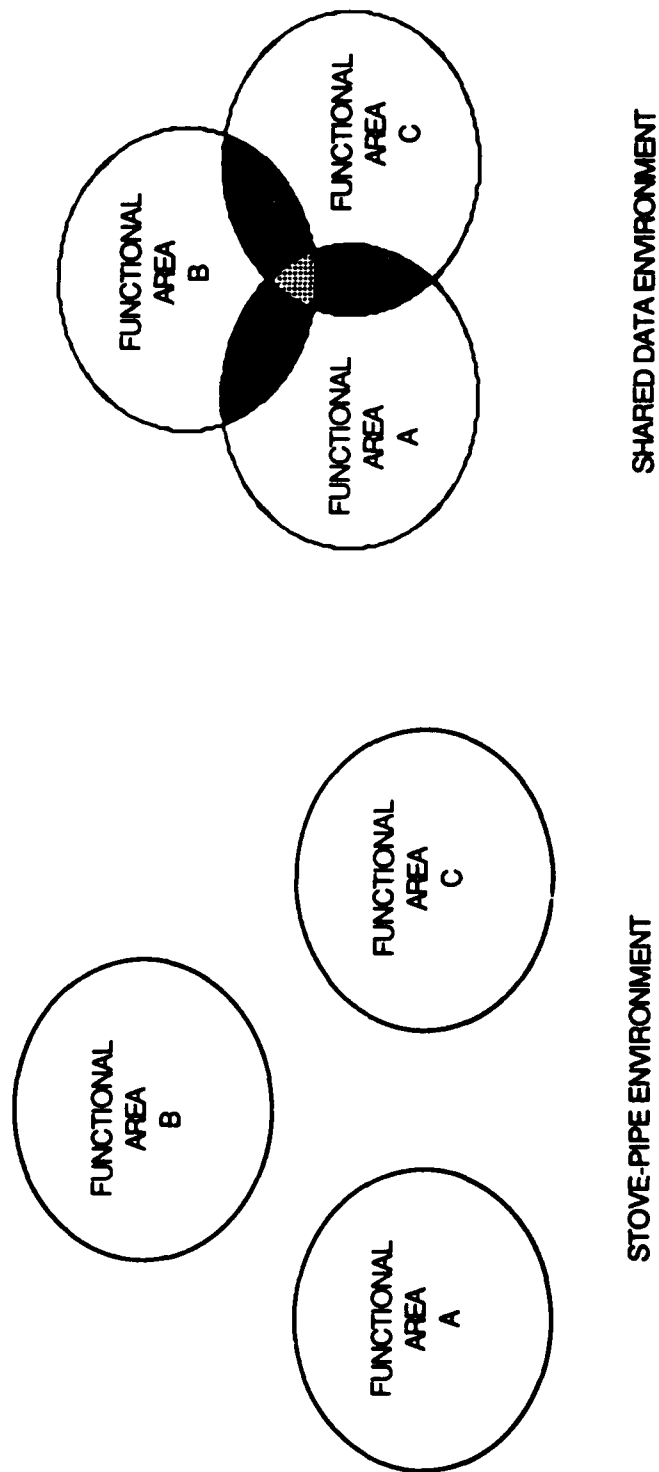


FIGURE 2-4
HYPOTHETICAL ARMY DATA DELIVERY SYSTEMS
FOR THREE FUNCTIONAL AREAS

2.2.4.3 Design Analysis

Having identified the alternatives, the next step is to identify how the analysis will be performed and in particular, the type of cost benefit analysis to be performed, how benefits will be measured, the costs to be included, sunk costs and benefits, the salvage values, the time period under analysis, and the discount rate to be used. Although each of these are important, the type of analysis, the costs to be included and the measurement of benefits are the more critical. These are discussed below.

In the TAPDB cost benefit analysis, the type of analysis was incremental. It set out to determine the incremental cost and benefits of ways to share personnel information among the components over a baseline approach, i.e., transmitting data via tapes.

The costs of both alternative "information delivery" systems is fairly straightforward, but depends upon the configuration assumptions used. As shown in the TAPDB study, the remote access method was the least costly. Similarly, a shared data approach which involves integrating different hardware and software (that is, a retrofit of an existing system) would be far more costly than one in which homogeneous hardware and software are purchased off-the-shelf.

A suitable measurement of benefits (system and user) is the most difficult and critical element of the analysis. For an Army data environment, decision-makers need information to make decisions relative to their area of responsibility. It is important to quantify, in some fashion, the value of the data being acted upon. For the two alternatives, one needs to determine the benefits of the data being collected, maintained, processed, and reported. For both alternatives, the question is: what is the combined value to the using community of the quality of data being received? From our literature search, we have reached the conclusion that there is no simple metric for making user benefits transparent and calculable (however, it can probably be effectively modelled.) A differentiation of user perceptions as to value received in shared and non-shared environments can only be accomplished through user surveys, Delphi techniques, or other forms of direct interaction with the users to identify and quantify the system's effect on the decision-making process and ultimate decision outcome. The objective of these direct interactions should be to doggedly pursue dollar quantification of these benefits; this permits a more unambiguous interpretation of the results of the analysis.

2.2.4.4 The Cost Benefit Report Format

The results of the cost-benefit analysis are generally presented in terms of "net present value" (NPV). This means that all the benefits of an alternative over an assumed life are treated as positive flow, and all costs

of an alternative are treated as negative. The NPV represents the discounted value of the benefits and cost. Guidelines are presented in OMB's Circular A-11. Typically, the alternative with the higher NPV is the preferred alternative. Benefit-to-cost ratios can also be used. Table 2-3 displays a proposed report format of the NPV of costs and benefits for the stove-pipe and shared data alternatives. As discussed in section 2.2.3, MITRE advocates through the ILCA approach, splitting the analysis into separate system efficiency and user effectiveness calculations. This provides visibility to the two aspects, and also allows for the possibility that the user effectiveness analysis could be avoided, or, at least, highly simplified. This could happen if the least cost system could qualitatively be expected to have user benefits at least the equal of the higher cost alternative. Therefore, dominance could be established on the basis of a mix of quantitative and qualitative results.

The costs and benefits shown in Table 2-3 are only representative of the types of elements that should be identified and evaluated. The actual cost breakdown structure and benefit listing should be selected on the basis of a specific understanding of the particular environment and system.

Although the conceptual framework can be described in general terms, the ultimate test of its usefulness can only be determined through specific application in a selected Army information setting. In fact, the only successful applications of user perceptions and benefit valuations found in the literature search centered around specific corporate information systems (e.g., Matlin, 1978). Additional support for this kind of test and evaluation framework has been received from several commands--e.g., PERSINCOM (Hollist, 1989 in Appendix B) and the Navy Personnel Research and Development Center (NPRDC) (Silverman, 1989 in Appendix B). At the outset on this task, military personnel systems had been singled out by ISEC data management specialists (Glymph, 1988 in Appendix B) as a potentially fruitful area for analysis.

Although the most logical "test bed" for the conceptual framework might prove to be a Standard Army Management Information System (STAMIS) program in a non-personnel area, selection and evaluation of an Army personnel information system (e.g., TAPDB) could have these advantages:

- Exploits opportunity for synergism in working with a group of information specialists currently active and experienced in the area of data redundancy and willing to cooperate in the benefit valuation process (Hollist, 1989 in Appendix B).
- Deals with a functional area (promotions/duty assignments, etc.) where decision processes and outcomes to be evaluated are more tractable.

TABLE 2-3
PROPOSED REPORT FORMAT OF NPV COSTS AND BENEFITS FOR
ARMY STOVE-PIPE VERSUS SHARED DATA ENVIRONMENTS

<u>COSTS</u>	(Thousands of FY89 Dollars)	
	<u>STOVE-PIPE</u>	<u>SHARED</u>
Facilities	\$	\$
Hardware		
Processors		
Communications		
Peripherals		
Work Stations		
Software		
Development		
System Concept		
System/Software Requirements		
Preliminary Design		
Detail Design		
Code/Test		
System T & E		
Operational T & E		
Quality Assurance		
Documentation		
Maintenance		
Facilities		
Hardware		
Software		
Operations		
Personnel		
ADP Supplies		
System Life Cycle Costs	\$ A	\$ B
<u>User Benefits</u>		
Net Effect on Decision-making Process		
- Quality of Output Contents		
- Quality of Output Form		
- Timeliness		
- Quality of System as an Aid to Problem Solving		
- System Stability		
- System Flexibility		
Impact on Decision Outcome		
- Value of Objective 1		
- Value of Objective n		
Value of Data to Decision-makers	\$ C	\$ D
Net Present Value	\$C-A	\$D-B

- Complements a parallel effort to investigate metrics for personnel information system users (Silverman, 1989 in Appendix B).

SECTION 3

CONCLUSIONS AND RECOMMENDATIONS

This section highlights the major conclusions of the literature search and discussions with information specialists on the data economics issue and specifically keys MITRE's findings to the objectives of this task:

- Survey of previous research,
- Assessment of data redundancy in both the commercial sector and within Army information systems, and
- Evaluation of approaches for measuring impact of data quality.

3.1 SURVEY OF PREVIOUS RESEARCH

MITRE performed a detailed literature search, reviewed over 100 documents, and prepared abstracts for 62 of the most relevant of these. Similarly, MITRE interviewed over 45 ADP and information specialists from the DoD, academia, and the corporate sector and prepared 32 interview reports of these conversations. Conclusions arising from these research efforts are as follows:

- Data redundancy is perceived as a significant problem from the perspective of both the system developer and system user. Data standardization and data dictionary efforts are underway in many military and corporate organizations to deal with the problem. [Army's PERSINCOM and Decision Support System (Hollist, 1989; Hoffman, 1988 in Appendix B), Navy's NPRDC (Silverman, 1989 in Appendix B), and Air Force Staff's Systems, C4 (Nguyen, 1989 in Appendix B)] This includes development and use of specialized data administration software to cope with many of the design and operational aspects of the problem. The effort has been a mixed success--ranging from significant reduction in redundant data elements to lengthy, ineffective efforts yielding minute progress (Ross, 1981).
- Data on information value to the user is a critical concept, often overlooked, that is just now receiving the attention it demands. The notion that an information system should be justified on the basis of its usage rather than solely by its efficiency has been surprisingly relegated to the "back burner" as "too hard". However, information specialists and academics have kept the issue alive with a considerable flow of "think piece" papers, theoretical approaches,

conceptual frameworks, and occasional applications. The federal government, with the Paperwork Reduction Act of 1980 and its official certification through IRM that information is not a free good but a resource that needs to be managed, has spurred further interest in information cost and value.

- No operations research methodologies were found in the literature search that appeared feasible for application to the information valuation problem. Although decision theory/game theory provide methodologies for incorporating many of the complexities in the quality of data versus quality of decision-making issue, the fact remains that the subjective probability distributions required by the analysis must be obtained by a degree of user interaction that is better accommodated in other analytical frameworks and approaches (e.g., ILCA).

3.2 ASSESSMENT OF DATA REDUNDANCY

Very little quantitative research has been done which definitively answers the question of whether a shared data environment is more cost-effective than non-shared data environments. Moreover, few documents were found which measured the value of processed data within a single information system. However, some research has been performed which grossly estimates the cost of "data redundancy" within a single information system. Other research measured the costs and benefits of an Army data sharing environment and a number of researchers have written about the various methods of valuing processed data (information) from efficiency and effectiveness viewpoints. The salient findings from our literature search are summarized below.

- Data redundancy (defined as a state in which data is defined and/or maintained by more than one source) is prevalent in many organizations' information systems and is roughly estimated at 75 percent.
- The cost of data redundancy within single corporate information systems has been roughly estimated at 50 percent of an organization's maintenance effort.
- Data redundancy and data inconsistency (defined as the same named data element appearing in two information systems which do not contain the same information or which possess different formats) is controllable with a strong Data Administrator, and forced use of automated data element dictionaries and CASE tools. Forced use of automated tools throughout the software life cycle creates data standardization and can substantially reduce software development and maintenance costs and moderate hardware requirements. However,

no research was found which rigorously measured these potential savings.

- Potential benefits of a shared data environment in the corporate setting were stated as:
 - Return on investment will be increased due to lower costs
 - Better information will permit the firm to attain a more competitive edge, maintain market leadership or enter a new product/market area
 - Software development cost will be shared
 - Maintenance costs will be shared
 - System reliability will be increased
 - Operating costs will be shared
 - Resource allocation decisions will be more efficient
- One researcher found that a wholesale distributor achieved a 30 percent return on the investment in a shared environment because of a 40 percent reduction in order/buying personnel and a reduction in processing equipment costs.
- In spite of such possible benefits, one researcher found that 25 percent of all real-time and telecommunications-based software projects fail, primarily because of the "bad data" and the difficulty of getting the meaning of data straight among organizational users.
- Within the Army environment, one study was found which measured the incremental costs and benefits of the various alternate ways to share personnel data among four personnel components. This analysis compared the costs and benefits of technically integrating four personnel systems over the baseline of merely sending data tapes. Results were presented in cost-to-benefit ratios and the best ratio was achieved by a remote access alternative. Benefits in this study were measured in data processing efficiency terms only. The value of obtaining better information or the resulting decisions was not addressed or quantified.

3.3 EVALUATION OF APPROACHES FOR MEASURING IMPACT OF DATA QUALITY ON INFORMATION SYSTEMS

The survey of approaches found in the literature search can be placed in the following typology:

- System Efficiency--Represented most prominently by cost-benefit analysis that concentrated on internal system cost savings, and either ignored user benefits or appended brief qualitative descriptions of their potential existence.
- User Effectiveness--Represented by complex decision and game theoretical approaches that focus exclusively on decision outcomes, essentially ignoring intermediate input and process effects.
- Synthesized Efficiency-Effectiveness--Characterized by an information life cycle approach that separately looks at the system's efficiency (i.e., similar to the cost-benefit approach) and user effectiveness (i.e., examines how he uses the information to reach a decision, and the outcome of that decision).

The synthesized approach, which MITRE calls information Life Cycle Approach (ILCA), is recommended as the approach most applicable and adaptable to Army information system evaluations. Although not a formal methodology, ILCA is a blend of the other two approaches that incorporates the essential elements of each, while providing the necessarily broader perspective required for assessing the impact of data/information. Moreover, it directly addresses the key issue of information system evaluation--its value in use--by specifically recognizing that how users use or value information in reaching decisions and the value of decision outcomes are uncertainties that can only be resolved through direct interaction and guided discussion with those users.

Finally, a strawman, conceptual ILCA framework--that sketches out some of the prospective system costs and benefits that would be necessary to evaluate the impact of data redundancy in stove-pipe and shared data environments--has been developed.

3.4 RECOMMENDATIONS

Since the literature search indicates that data/information value and impact analysis is an evolving area of great interest in the information community, a general recommendation is to closely monitor the major periodicals cited in this report for new developments and, more specifically, maintain a close watch on the current PERSINCOM and Naval Personnel R&D Center (NPRDC) efforts in data redundancy and user metrics, respectively.

Finally, MITRE recommends that the usefulness and feasibility of the ILCA approach be assessed in a specific application or demonstration project. A candidate information system could be selected from the PERSINCOM or STAMIS systems for evaluation, including a data collection effort centered on the identification and valuation of user perceptions.

APPENDIX A

DOCUMENT ABSTRACTS

Over 100 documents, reports, and periodical articles were reviewed in the course of MITRE's literature search. These data sources included DOD, Federal Government, industry, and academic references. Brief abstracts of 62 of these references are included in this Appendix, arranged alphabetically by author.

TITLE: A Systematic Approach Toward Assessing the Value of an
Information System

ORGANIZATION: University of British Columbia

AUTHOR: Ahituv, Niv

DATE: December 1980

REPORT NUMBER/PUBLISHER: MIS Quarterly, Vol.4, No.4

ABSTRACT:

The author classifies the various approaches toward information system evaluation into two categories: pragmatic assessment, such as cost-benefit analysis; and theoretical evaluation based on decision theory. This paper suggests a third approach, which attempts to marry the empirical and theoretical methods.

He begins by listing possible attributes of a user utility function associated with reporting systems. Those attributes selected include timeliness, contents, and format. The next step is defining a possible measure for the attribute (response time; similarity between ex ante and ex post data requirements; and subjective evaluation of report medium/data ordering/graphic design, respectively). The final steps--and the more complex ones--deal with assessment of the utility gained by the user from each individual attribute, and the combining of these utilities in a joint utility function. The author admits that the ideal situation for optimization in this example would require that every relevant attribute is known and measurable; that the utility function related to each individual attribute is clearly defined; and the tradeoff among the various attributes are available and provide a clear mathematical formulation of the joint utility function. Given the unlikelihood of the ideal, he closes with a non-optimal solution to an illustrative source selection problem that is clear and implementable.

TITLE: Information Issues in Model Selection

ORGANIZATION: Massachusetts Institute of Technology (MIT)
University of Texas at Austin (UT)

AUTHOR: Balakrishnan, Anantaram (MIT) and Whinston, Andrew B. (UT)

DATE: December 1988

REPORT NUMBER/PUBLISHER: National Science Foundation, Grant No. ECS-
87186734

ABSTRACT:

This paper focuses on an aspect of management science that has received relatively little attention in the operations research literature--the model building process, or more specifically, the process which guides the selection of an appropriate model for a given decision context. The authors explore one aspect of this modelling science, namely, the role of data acquisition and informational uncertainty in model building. They consider a class of decision problems where the modeler faces a wide spectrum of informational gathering choices that vary in the amount and quality of information that they provide as well as their costs and delays. They make the point that, if acquiring additional information (or improving the quality of your present information) has diminishing returns--in terms of the quality of final decisions suggested by the model--then partial information gathering or improvement might be the most appropriate or cost-effective strategy. This information-decision quality tradeoff is illustrated by the use of several production and traffic examples.

Throughout these examples, detailed and precise models are equated with a higher level of data acquisition and cost. The problems are therefore easily translatable to the problem of MIS informational costs and benefits.

The authors persuasively argue that the decision on how much and what type of information to collect and encode in the decision model is an early and critical part of the model-building exercise. The tradeoff analysis, however, assumes that adequate quantification of the payoffs to the decisionmaker for the various information gathering levels can be developed. The value of information remains as the most intractable feature of the process.

TITLE: Reconciliation Process for Data Management in Distributed Environments

ORGANIZATION: School of Business, State University of New York

AUTHOR: Ballou, Donald and Tayi, Giri Kumar

DATE: June 1985

REPORT NUMBER/PUBLISHER: MIS Quarterly

ABSTRACT:

The trend toward distribution of information resources throughout an organization requires an increased awareness of, and sensitivity to the organization's structure, strategy and constraints. However, organizational requirements may be in conflict with technical concerns involving distributed systems. The authors present a mechanism for analyzing the degree of consistency regarding MIS activities level of management.

The authors discuss three areas of consideration when attempting to decide exactly how a distributed environment should be implemented. These areas are strategic, managerial and operational.

Data integrity, responsibility, usage patterns, update characteristics, storage/transmission cost, data availability, and experience with DDP are factors which need to be addressed to determine how a DDP should be implemented.

TITLE: Inter-Organization Information Sharing Systems

ORGANIZATION: N/A

AUTHOR: Barrett, Stephanie and Konsynski, Benn

PUBLICATION/PUBLISHER: MIS Quarterly

DATE: Special Issue 1982

ABSTRACT:

The authors propose a classification scheme for inter-organization information sharing systems. The classification system was based upon a survey of commercially existing DDP implementations.

The authors present three basic incentives for moving to a shared data environment: cost reductions, productivity improvements, and enhanced product/ market position. The authors discuss how these incentives as well as complexities of an organization's operating environment determine the implementation approach.

In presenting the incentive for cost reductions, the authors cite one company within their study which achieved substantial cost saving. The company's return on investment was 30 percent and they reduced personnel expense by 40 percent. The authors also state that a shared environment decreases error rates and that leads to administrative cost avoidance. Other cost avoidance features in a shared environment are the elimination of data entry, the sharing of control costs, cost displacement, and increased productivity. The authors clearly state that the degree of cost reductions are dependent on the size and scope of the shared information system, the number and type of participants, and the chosen implementation approach.

TITLE: Economic Theory and Operations Analysis

ORGANIZATION: Princeton University

AUTHOR: Baumol, William J.

REPORT NUMBER/PUBLISHER: Prentice-Hall, Inc.

DATE: 1965

ABSTRACT:

Baumol has been at the forefront of those economists advocating closer links between theory and business practices. In this book, he uses operations research to bridge the two streams of thought, in the process providing excellent overviews of microeconomic and mathematical economic analysis and concepts, which can be usefully applied to solving business problems. There is a separate chapter devoted to decision theory, which contains clear and concise distinctions between utility, game, and decision theory, followed by discussions and illustrations of decision rules (e.g., maximin, maximax, Bayes/Laplace, and minimum regret criteria). These are short, worthwhile descriptions to grasp, but the book offers little else of direct relevance to the ISC task.

TITLE: Quantitative Analysis for Business Decisions

ORGANIZATION: Cornell University, Stanford University

AUTHOR: Bierman, Harold, Jr.; Bonini, Charles P.; and Hausman, Warren H.

DATE: 1977

REPORT NUMBER/PUBLISHER: Richard D. Irwin, Inc.

ABSTRACT:

This is a book about business decision-making. The purpose is to describe a representative sample of the models (and their related quantitative techniques) that are frequently used in the process of evaluating and selecting the various courses of action open to the decision-maker. A major portion of the book is devoted to decision-making theory and its relationship to classical statistics, probability distribution, and utility theory. The authors make an interesting distinction between "classical statistics" with its primary focus on scientific or experimental work where the probability of error is more accurately measured than the consequences of a decision, and "business decision-making", where objective probabilities are lacking, but consequences are very important and subject to measurement. The book offers a very good linear progression from a survey of quantitative analysis and probability concepts through mathematical programming and deterministic and probabilistic models. Its primary value to the ISC task is its excellent step-by-step review of the decision process, and the subjective elements required to achieve the expected value maximization criterion. Objectives and consequences, measurement techniques, and the choice of probabilities for possible states of nature (e.g., data redundancy) are, to some extent, judgmental.

TITLE: Assessing the Value of Information in Organizations: A Challenge for the 1980s

ORGANIZATION: Bureau of Govt Research and Service
University of South Carolina

AUTHOR: Black, Sena H. and Marchand, Donald A.

DATE: 1982

REPORT NUMBER/PUBLISHER: The Information Society Journal, Vol. 1, No.3

ABSTRACT:

This paper addresses the issue of increasing resources in the development and use of information services and technologies without a concomitant effort to assess the value of such information. The purpose of the paper is to evaluate some major perspectives or ways of assessing information value and to discuss their relevance to the task of understanding information use in organizational settings. This evaluation is conducted by isolating three major dimensions for systematic comparison: quantity versus quality of information focus; functional versus process orientation; and value criteria selection.

The survey of alternative approaches for assessing information value include the following:

- Information Technology
- Information/Library Science
- Information Resource Management
- Value Burden
- Organizational Theory
- Economics of Information

The authors discuss each of these approaches in detail, outlining key assumptions, objectives, inferences, unit of analysis for information interpretations, and criteria of value assessment. These are all summarized in a most useful and informative overview table. In an interesting aside, these approaches are characterized by the authors as representing a continuum of quantity versus quality concerns, with Information Technology and Library Science representing the quantity focus while the Organizational Theory and Economics approaches concentrate on quality of information.

The paper concludes with a proposed conceptual framework for information value assessment in organizational settings which synthesizes the "best" features of the previous approaches.

TITLE: Infomap: A Complete Guide to Discovering Corporate Information Sources

ORGANIZATION: N/A

AUTHOR: Burk, Cornelius F., Jr. and Horton, Forest W., Jr.

DATE: 1988

REPORT NUMBER/PUBLISHER: Prentice-Hall

ABSTRACT:

Infomap is a guidebook to information designed for the use of the senior corporation or activity officer. Its unusual title derives from its exploratory mining analogy. The authors' purpose is to assist the enterprise to discover, map, and evaluate its information resources (the "gold mine" of this information technology age). Although long on descriptions and analogies and short on innovative methodological development, it does synthesize previous studies effectively and makes use of their results in this cookbook approach to constructing an evaluative framework for information. Of primary interest to this ISC task is the chapter on "Measuring Costs and Assessing Values". Although their exploration of costs is somewhat simplistic, the identification of MIS cost elements is worthwhile. Similarly, value elements in information systems are identified and classified as follows:

- Quality of Information Itself (accuracy, currency)
- Utility of Information (ease of use, format)
- Impact on Organizational Productivity (improvements in decision-making)
- Impact on Organizational Effectiveness (meeting goals and objectives)
- Impact on Financial Position (cost reduction/savings, return on investment)

As hinted in the title of the chapter, the authors are pessimistic in "measuring" value and settle for assessing, which is heavily subjective and qualitative. Nonetheless, they do develop a "cost-effectiveness" framework where costs and benefits--although presented in different units--are rank ordered; and cost/value indexes are used to make cost-effective choices.

Because the book was recently published, it also contains an excellent up-to-date bibliography.

TITLE: Where Methodology Falls Short

ORGANIZATION: N/A

AUTHOR: Carlyle, Ralph

DATE: December 1, 1988

REPORT NUMBER/PUBLISHER: Datamation

ABSTRACT:

The author discusses a number of reasons why 25 to 40 percent of all real-time and telecommunications projects are never completed. The author argues that the biggest reason for these failures is due to the "lack of a review structure that allows (management) to monitor the progress of an application and input changing plans and strategies." The author presses for complete automation of the software development and maintenance process, i.e., Computer Aided Software Engineering (CASE) tools.

The author cites several cases in the power industry where databases were unnormalized and data inconsistencies have lead to rate increases or adjustments to the utilities. Interestingly, the author notes that moving to a shared data environment alone will not guarantee data consistency--the problem can become more acute. Without an automated tool which forces discipline in the creation of new applications (and hence data), "...programmers frequently feel too rushed to do a thorough job with the logical data models and begin to backslide into old ways."

TITLE: IRM: A Look Backward and Forward at the Federal Level

ORGANIZATION: Auburn University

AUTHOR: Caudle, Sharon L.

DATE: Spring 1988

REPORT NUMBER/PUBLISHER: Information Management Review, Vol. 3, No. 4

ABSTRACT:

This paper reports on the findings of an empirical study of Information Resources Management (IRM) implementation experiences. IRM was introduced by the Paperwork Reduction Act of 1980 as a management concept to address the many problems then existing in federal reporting requirements and technology integration and applications. It was defined as "the planning, budgetary, organizing, directing, training, promoting, controlling, and other managerial activities involved with the collection or creation, use, and dissemination of information by Federal departments and agencies." The most powerful requirements of the Act were the designation of senior IRM officials in each agency, oversight of information collection, triennial reviews of information management activities (especially acquisitions and the use of information technology), and five year information technology plans. Caudle, who is a recognized expert in IRM and who has served with OMB and the National Academy of Public Administration, concludes that IRM implementation has proceeded much more slowly than first projected. In fact, there has been a basic clash in objectives. One perspective of IRM is program service and responsiveness, while the other stresses oversight and administrative needs. Improvement in performance seems to have been sacrificed to the addition of another level of bureaucratic inertia. There are some exceptions to the lack of progress which Caudle does not point out--notably, efforts to manage the information resource in a more cost-effective manner (Chick, 1984).

TITLE: A Multiple Criteria Approach for Evaluating Information Systems

ORGANIZATION: University of Illinois, Department of Accountancy

AUTHOR: Chandler, John S.

DATE: March, 1982

REPORT NUMBER/PUBLISHER: MIS Quarterly, Vol. 6, No.1

ABSTRACT:

The approach described in this paper provides a framework for gaining insight into information system performance from both the user and system viewpoints by establishing a causal relationship between user goal attainment and system activity. The underlying principle is that an information system is a symbiotic relationship between the users and the system itself, and that equal consideration has to be given to both user and system constraints. Therefore, performance has to be measured in terms of the computer system domain (resource utilization, cost, efficiency) and the user domain (throughput, reliability, response time). The complexity of the evaluation problem has increased significantly for modern computer based information systems because of an expanding range of users and applications with a corresponding expansion of diverse performance goals and resource requirements, and a growing demand to achieve conflicting performance objectives (e.g. time versus cost versus effectiveness).

The author's approach is based on a multiple goal programming formulation of the information system design which is designed to minimize the discrepancy between user goal expectations and actual system performance for all goals. The formulation, because it deals with multi-dimensional situations, is quite complex. Due to internally conflicting goals, the method does not produce an optimal solution--only a feasible, satisfactory one that minimizes total discrepancies. Finally, the issue of goal quantification is only raised; no solutions are proposed.

TITLE: Decision Analysis for Industry

ORGANIZATION: Stanford Research Institute

AUTHOR: Chen, Kan

DATE: 1968

REPORT NUMBER/PUBLISHER: Unpublished Paper

ABSTRACT:

The author discusses decision analysis procedures in an industrial management setting by reviewing three phases of analysis:

- Deterministic--how much should be allocated to conduct study (the author uses a one percent rule against the total value of the decision); identify alternatives; assign values to outcomes (profit); select major variables and establish relationships between/among variables; determine dominance and sensitivity of solutions.
- Probabilistic--encode uncertainty on variables (subjective probability/profit lottery); encode risk preference (concept of utility); selection of alternative with highest expected utility.
- Post Mortem--develops value of perfect information and the most economical program of additional information gathering. Purpose of this phase is to identify what information is relevant to the decision and how much the additional information is worth.

The paper presents some interesting insights, but these are explained in more detail in other references.

TITLE: Information Value and Cost Measures for Use as Management Tools

ORGANIZATION: General Accounting Office (GAO)

AUTHOR: Chick, Morey J.

REPORT NUMBER/PUBLISHER: National Bureau of Standards (NBSIR 86-3324)

DATE: February 1986

ABSTRACT:

This article presents some ideas on approaches for establishing and measuring the value and cost of information and how this analysis can be used as a management tool in IRM. The author claims that the concepts of information value and information costs are not fully comprehended. Although research has been performed in these areas, progress has been hampered by the lack of accounting methodologies for tracking information costs; by the intangible nature of information and the need to allocate any cost among its users and various applications; and by the failure to reach consensus on the notion of information value--especially the need for assigning monetary measures. Chick presents a detailed schematic describing the resources applied to producing information; this is presented as a guide to the type of accounting approach needed to capture information cost. Furthermore, he provides some possible indicators of information value as well as appraising their potential for quantification. These include: impact on income; value-added motivations; reduction in costs from use of the information; impact of information problems (e.g., redundancy); usefulness and impact of information use on defined organizational goals (e.g., effectiveness); productivity and efficiency improvements; and user perceptions of value. There is a short, but interesting, discussion of these indicators followed by a detailed chart which summarizes the cost and value impact of various categories of information problems. For example, data quality and duplication problems are described in terms of their impact on cost and value, and examples from recent Federal experiences are cited.

Chick, both in his articles and in discussions, provides some of the best insights available on the issue of information value.

TITLE: Policies and Research Needed for Assessing Risk of Automated Information Used in Human Safety Applications

ORGANIZATION: GAO

AUTHOR: Chick, Morey J.

DATE: Summer 1986

PUBLICATION/PUBLISHER: Information Management Review, Vol. 2, No. 1

ABSTRACT:

Many automated information system applications exist that can affect the public's health and safety. Failures of these system applications could and have put human lives in jeopardy (e.g., diagnosis of medical patients, monitoring of nuclear plants, air traffic control, etc.). Chick examines these systems, reviews some past failures, and reviews the conditions, which are likely to lead to the failures. The causes of failure are categorized as data quality/availability, software development and maintenance, security, information management/people, and hardware. No attempt is made to quantify the effects of these failures although he mentions airline crashes, the near-disaster of Three Mile Island, and other examples where erroneous decisions were made and actions taken either by automated systems or by people heavily reliant on those systems. In the face of the growing threat, the author calls for a sharper, centralized focus on the problems, conditions, and causes of automated information systems failures.

TITLE: Assessing the Impact of Information Technology on Enterprise Level Performance

ORGANIZATION: Center for Information Systems Research, MIT

AUTHOR: Crowston, Kevin and Treacy, Michael E.

DATE: October 1986

REPORT NUMBER/PUBLISHER: CISR Working Paper No.143

ABSTRACT:

In this paper, the authors review research that has been performed on enterprise level impacts of information systems, with a particular emphasis on research that has attempted to measure those impacts. The authors studied eleven articles relating to enterprise level performance; these represent methodologies ranging from cost-benefit analyses and field surveys to conceptual frameworks. Most of the cost-benefit studies are based on a simple accounting view of productivity and do not address the issue of improved managerial performance. The conceptual framework approaches are drawn from microeconomics, specifically focussing on technical efficiency and efficient frontier analysis. The authors conclude that all the methodologies fall short in defining and measuring firm inputs and outputs affected by the information technology. Moreover, those studies that do make the effort to identify and measure specific input and output variables lack general applicability and validity. The major shortcoming, in the view of the authors, is the failure to understand or frame a theory of the information processes within the firm. Understanding this can lead to a clearer choice of variables and to the generation of testable hypotheses about the impact of information on overall enterprise productivity.

TITLE: An Analysis of the Impact of Distributed Data Processing on
Organizations in the 1980's

ORGANIZATION: N/A

AUTHOR: Davis, Charles and Wetherbe, James

DATE: December 1979

REPORT NUMBER/PUBLISHER: MIS Quarterly

ABSTRACT:

The authors examine "Distributed Data Processing (DDP)" by defining it, projecting its impact on organizations, and assessing the economic aspects of the concept.

The authors state there is much confusion as to exactly what DDP is. The term DDP has been used to refer to highly centralized computing systems with remote terminals at one extreme, to totally decentralized collections of stand alone minicomputer systems on the other. The authors offer a framework for considering DDP. This framework is a continuum with centralized computing at one extreme and decentralized at the other.

Very little information is available about data integrity, although they state that distributed database systems are technically difficult to develop; yet the demand is high.

TITLE: Rational Choice in an Uncertain World

ORGANIZATION: Carnegie-Mellon University

AUTHOR: Dawes, Robyn M.

DATE: 1988

REPORT NUMBER/PUBLISHER: Harcourt, Brace, and Jovanovich, Inc.

ABSTRACT:

Dawes, in an article called "Proper and Improper Linear Models," asserts that random, unit weights placed in a linear model consistently outperform intuitive, global judgments of trained experts. Because decision-makers-- without the discipline of a model--will allow inconsistent subjective judgments to enter into the analysis, models with arbitrarily, but reasonably weighted variables perform better. He says that there is a bias of feedback that gives intuition a seemingly higher grade--when intuitions perform well, one hears about it. If they fail, no one does. In contrast, systematic predictions of linear models yield data on how poorly they predict (variance, etc.).

Some of these descriptive arguments are relevant to the issue of utility functions in multiattribute decision analysis.

TITLE: Information Planning Guidance

ORGANIZATION: USA DISC4

AUTHOR: DISC4

DATE: January 29, 1988

REPORT NUMBER/PUBLISHER: DISC4

ABSTRACT:

This document provides the top level guidance for the development of information management plans by functional proponents within the Army. Many data information systems plans are described.

TITLE: Decision Analysis in Clinical Patient Management with an Application to the Pleural Effusion Syndrome

ORGANIZATION: RAND Corporation

AUTHOR: Ginsberg, A. S.

DATE: July 1971

REPORT NUMBER/PUBLISHER: R-751-RC/NLM

ABSTRACT:

This report proposes and illustrates a practical analytical decision-making aid for the solution of typical patient-management problems. The problem is defined as choosing a course of action to follow with a patient exhibiting a particular set of syndromes. The course of action is chosen from a set of alternative acts (e.g., performing diagnostic tests, administering therapies, waiting) with the objective of achieving the best possible results from his and the patient's point of view. The approach relies heavily on the concepts of decision analysis, and considers the uncertainties and risks inherent in the decision problem (such as the dollar costs and risks of complications associated with a diagnostic test or treatment). A simple hypothetical example, which illustrates the use of decision analysis and the maximization of expected utility, explains the basic ideas. The model is remarkable in its detailed approach to monetizing the consequences/complications, such as: out-of-pocket dollar costs (tests, care, lost income); days in bed with no, mild, or severe pain; days for diagnosis or definite action; permanent or temporary complications; short-term severe pain; long-term mild pain; days of restricted activity; death. Rather than develop dollar tradeoffs for each complication, Ginsberg uses subject and expert interviews and a lottery methodology to make relative comparisons among complications (i.e., decreasing preference), thereafter assigning dollar values for his utility function.

TITLE: The Value of Information and Related Systems, Products, and Services

ORGANIZATION: King Research, Inc.

AUTHOR: Griffiths, Jose-Marie

DATE: 1982

PUBLICATION/PUBLISHER: Annual Review of Information Science and Technology
(ARIST)

ABSTRACT:

The author reviews the literature available on the issue of value in information systems by examining the general, economic concepts of value. It is described as an attribute that establishes an equivalence relationship (i.e., in exchange). Three further characteristics are discussed: its subjective nature, the situation-dependency of its assessment, and its variance over time. He defines a total value function made up of three user perspectives:

- Willingness to pay, or "input" perspective. This, he calls, apparent value.
- How information affects work ("process" perspective). This is also termed consequential value and is measured by the average savings in terms of time and materials attributable to the use of the information.
- How the work above affects the environment or organization ("output" perspective). This higher order consequential value can be measured in several ways (e.g., return on investment).

He notes that the valuation becomes more difficult as you move from the input to the output perspective.

One of the studies cited describes the relative value of information systems with these dimensions:

- Data or information quality (precision, accuracy, source credibility)
- Scope (recency and completeness)
- Hassle (time lag, ease of access, ease of use)
- Cost

Although the importance of value assessments is growing in the face of budget cuts and economic restraints, progress in methodological development has lagged. It is hoped that some of the conceptual frameworks cited in the paper will be expanded.

TITLE: Evaluating Information System Effectiveness--Part I: Comparing Evaluation Approaches

ORGANIZATION: School of Management, University of Minnesota

AUTHOR: Hamilton, Scott and Chervany, Norman L.

DATE: September 1981

REPORT NUMBER/PUBLICATION: MIS Quarterly, Vol 5, No.3

ABSTRACT:

The purpose of this paper is to provide an overview of approaches for evaluating MIS effectiveness. This is complicated by the system's multidimensionality, qualitative and quantitative aspects, and multiple--and often conflicting--evaluator viewpoints. System effectiveness is addressed through a conceptual hierarchy of system objectives differentiated between efficiency-oriented (i.e., the MIS process) and the effectiveness-oriented (which focuses on the user process and performance). Assessments of system effectiveness in meaningful terms are frequently hampered by system objectives and performance measures which have been inadequately defined, which tend to be efficiency-oriented and easily quantified, and which continually evolve. The author recommends measures for improved performance in these areas, and sketches out several approaches that can be used, including:

- Service level monitoring--assessment of value to the user based on the terms established between MIS and user personnel (e.g., turn-around time).
- User attitude survey--questionnaires and interviews focusing on the user's perceptions (e.g., quality of reports).
- Post installation review--requirements definition (e.g., does the system do what it is designed to do?)
- Cost/Benefit analysis--quantifies the system's effect on organizational performance in terms of dollars (e.g., direct cost savings, tangible financial benefits).

The author makes the closing point that user perceptions, expert judgments, and other subjective means cannot be avoided in the evaluation process.

TITLE: Evaluating Information System Effectiveness--Part II: Comparing Evaluator Viewpoints

ORGANIZATION: School of Management, University of Minnesota

AUTHOR: Hamilton, Scott and Chervany, Norman L.

DATE: December 1981

REPORT NUMBER/PUBLISHER: MIS Quarterly, Vol.5, No.4

ABSTRACT:

The purpose of this paper is to examine the nature of the differences in assessments of system effectiveness between decision-makers in various functional groups involved in MIS implementation--users, MIS, internal audit, and management. The functional groups are first described, then the viewpoints are compared and contrasted on a pairwise basis. Several studies are referenced which address these viewpoints and rank order criteria of quality (e.g., user-identified accuracy, reliability, timeliness, assistance, and percentage of system objectives met). Finally, suggestions are made for incorporating multiple evaluator viewpoints into evaluation approaches. The most common procedures, for accounting for and resolving differences, involves participatory approaches, such as Delphi.

TITLE: Data Standardization Program

ORGANIZATION: The MITRE Corporation

AUTHOR: Herden, Melody

DATE: September 1987

REPORT NUMBER/PUBLISHER: MITRE, MTR-87W150

ABSTRACT:

This document identifies the major military and commercial standardization programs and associated organizations currently developing data element standards at the DOD-level and above. The author relates these efforts to the Army's data standardization efforts, presents findings and makes a number of recommendations about what the Army should do to improve its data elements.

In short, the author found and recommended:

- Data standardization and data management issues are neither centralized nor automated, and they should be.
- Standard data elements and codes from external programs are not being used to the extent they should be.
- The Army should continue and expand its participation in the working groups and committees of Army-external standardization organizations.
- Data elements are the core, fundamental resource of information resources management. Leadership for global information management issues is not centralized at the DOD level and should be. The author suggests the Army take this lead.

The report presents a solid rationale for data standardization and provides good background material concerning DOD's efforts to standardize data elements.

TITLE: Rational Choice: The Contract Between Economics and Psychology

ORGANIZATION: University of Chicago

AUTHOR: Hogarth, Robin M. and Reder, Melvin W.

DATE: 1987

REPORT NUMBER/PUBLISHER: University of Chicago Press

ABSTRACT:

This book brings in sharper focus the different perspectives of economists and psychologists, and their respective assumptions about rationality in decision-making. The violations of rational behavior have a decided impact on applied economics where rational economic behavior is assumed. The book contains the proceedings of a conference of economists and psychologists held to discuss this issue in the context of "do decision-makers act predictably to maximize their expected utility?" Economists have a different stake in this question since they are attempting to model "industry-level" behavior, whereas the psychologists have more of a traditional interest in individual differences; and therefore, rationality violations are more easily accommodated. Behavior--"rational" or otherwise--has to be viewed in a specific situation, and has to consider the goals it is attempting to realize, and the computational means available for determining how well the goals are reached. To understand behavior in decision-making requires observing processes directly while they are going on, either in real world situations or in the laboratory, and/or interrogating the decision-maker about beliefs, expectations, and methods of calculating and reasoning. In situations that are complex and in which information is very incomplete (i.e., "real world" situations), behavior even approximating an objective maximization of utilities is highly unlikely. What is more likely is a "satisficing" decision --a boundedly rational or reasonable decision. These concepts have relevance to any utility approach for the ISC data economics task, and reiterate the importance of the specific over the generic solution.

TITLE: Benefit Cost Analysis--A Delphi Approach

ORGANIZATION: Planning Research Corporation

AUTHOR: Horton, Forest Woody, Jr. and Pruden, John S.

DATE: Spring 1988

REPORT NUMBER/PUBLISHER: Information Management Review, Vol. 3, No. 4

ABSTRACT:

This paper focuses on the benefit-cost evaluation techniques used to appraise the Department of State's Foreign Affairs Information System (FAIS). The authors say that since the Department of State (DOS) only product is information, and since its effective use is the only measure of its success, FAIS is designed to be instrumental in collecting, processing, and disseminating field data and other key sources to decision-makers. FAIS must also meet the budget criteria set forth in OMB Circular A-11, namely the benefit-cost criteria. Because of its complexity in linkages and integration, the perceived benefits of FAIS could not--like costs--be tracked in a straightforward, conventional methodology. The department finally settled on the Delphi Technique, or the pooling and convergence of expert opinion. The course chosen was to compare the value enhancements expected for two alternatives at each of the nine stages of the information life cycle (i.e., draft through distribution, disposition, and administration). The benefits were estimated by following an algorithm with five parameters involving different variables including progress of utilization, percentage of the total users involved at each stage, quantified benefit based on salary and overhead factors, availability of functions factors, and a sensitivity analysis. Value enhancements were defined as time savings expressed as dollars spread across the entire organization and designed to capture the following kinds of illustrative benefits: improved productivity, improved quality of decision-making, improved performance of tasks, improved learning curve, faster response time, greater reuse of information assets, reduced reliance on paper files, etc. Non-recurring benefits, cost avoidances, and cost reductions were also considered. The results were then passed through the net present value calculations required by OMB Circular A-11 as well as validated by sampling of foreign service personnel.

TITLE: AR 25-1: The Army Information Management Program

ORGANIZATION: HQDA

AUTHOR: HQDA

DATE: March 1, 1986

REPORT NUMBER/PUBLISHER: HQDA

ABSTRACT:

This regulation establishes a methodology within the Army in which all plans and information systems are managed, resourced, and executed. This methodology is called the Army Information Management Program. The regulation establishes policy for program execution.

Briefly stated, the regulation provides a formal process for the development of new information systems. Each organization desiring to develop an information system must develop a logical data model, and architectural and system development plans. These documents must be reviewed by higher level to preclude development of unnecessary or redundant information systems and to encourage the sharing of information resources. The data model depicts "what" information requirements are needed in support of an organization's mission. The information architecture depicts "how" the information requirements will be satisfied. The system development plan describes who will implement the system.

A proposed initiative information model must be compared against the existing baseline of "information", and if a real need is not met, the requirement is made part of the information plan for development. This plan is revised each year to reflect changes in guidance and priorities. Once the initiative is included in the plan, the architectural plan (the how) will be assessed by higher levels to ensure interoperability with existing facilities, and is in keeping with technological advances. Once approved, specific organizations are responsible for developing the system to meet the requirement.

The regulation groups Army information into three categories and assigns responsible organizations for development and submission of information management plans for each of these groups. The three information groups are strategic, theater/tactical and sustaining base. These plans are made into a Information Management Master Plan.

TITLE: Operations Research in Decision Making

ORGANIZATION: Pennsylvania State University, U.S. Postal Service

AUTHOR: Ignizio, James P. and Gupta, Jatinger N. D.

DATE: 1975

REPORT NUMBER/PUBLISHER: Crane, Russak, and Company, Inc.

ABSTRACT:

This textbook, written to acquaint managers with the potential use of operations research (OR), is relatively easy to understand. Mathematical requirements are minimal, and all techniques are illustrated by examples. The book begins with an introductory chapter giving the definitions, history, and survey of common operations research models and approaches for decision-making. The rest of the book builds on that and focuses on the quantitative methods of OR as applied in: Decision Theory, Linear Programming, Network Analysis, Probability Models and Simulation, Inventory Systems, Queuing Models, Sequencing and Scheduling Problems. The Decision Theory chapter provides a framework for the general analysis of decisions, whereas the other chapters deal with decision methods for more specific classes of problems (e.g., linear programming, networks, etc.). The authors define three distinct classes of decision models: decisions under certainty (i.e., one possible outcome), decisions under risk (more than one possible outcome, but including meaningful probability distributions for those outcomes), and decisions under uncertainty (multiple outcomes, but without any meaningful probability distributions). Principles of expected value, most probable future, and aspiration level are used to achieve reasonable solutions under risk. Other strategies (minimax, maximin, Laplace's equal probabilities) are demonstrated under uncertainty.

Overall, the book provides a useful primer for these OR techniques, but does not furnish any directly applicable guidelines for the ISC task.

TITLE: Information Resources and Economic Productivity

ORGANIZATION: Center for Information Systems Research, MIT

AUTHOR: Jonscher, Charles

DATE: December 1983

REPORT NUMBER/PUBLISHER: CISR Working Paper # 110

ABSTRACT:

Jonscher uses a macroeconomic model of the U.S. economy to estimate the effect of information inputs on economic output and productivity. The analysis is based on a model of the interrelationship between the size of the information sector and the productivity performance of an economy. The theory is built up on the basis of a set of micro-analytic models of the effect of information inputs on the output or value of production and trading activities; it is then calibrated and applied for predictive purposes at the economy-wide level of aggregation. The purposes of the model are to explain the past growth of the information sector workforce, to identify productivity trends in the sector, and to determine the implications of those trends for future economic performance. As such, the methodology does not provide any useful insights for corporate-level impact evaluations, nor does the structure and complex formulations lend themselves to easy adaptation to the present ISC scenario.

TITLE: Value Analysis: Justifying Decision Support Systems

ORGANIZATION: Sloan School of Management, MIT

AUTHOR: Keen, Peter G. W.

DATE: March 1981

REPORT NUMBER/PUBLISHER: MIS Quarterly, Vol.5, No.1

ABSTRACT:

This paper is the author's appraisal of various evaluation techniques for assessing Decision Support System (DSS) proposals. He examined cost-benefit analysis, scoring techniques, and feasibility analysis. He finds that all require fairly precise estimates of, and tradeoffs between, costs and benefits, and often do not handle the qualitative issues central to DSS development. He notes that, even when cost-benefit analyses are formally required, they are avoided or used infrequently. Justification of system acquisition is more often based on mandated requirements, or alternatively, on identification of one or two benefits (rarely quantified).

Since innovations are value-driven, and managers seem to be more comfortable thinking in terms of perceived value and then asking if the cost is reasonable, he proposes use of "value analysis" in these evaluations. The complex calculations of cost-benefit analysis are replaced in value analysis by a two-stage set of simple questions. During the first (prototype) stage, he asks:

- What exactly will I get from the system?
- If the prototype costs \$x, do I feel that the cost is acceptable?

The key point is that value is considered first, and that costs and value are kept separate. The second stage considers full development, and asks these questions:

- How much will the full system cost?
- What threshold of values must be obtained to justify the cost?
- What is the likelihood that these values will occur?

If the expected values exceed the threshold, no further quantification is required. If they do not, then there must either be a scaling down of the system and a reduction in cost, or a more detailed exploration of benefits. The advantage of this approach is that, despite its focus on value, it does not require precision.

TITLE: Decisions With Multiple Objectives: Preferences and Value Tradeoffs

ORGANIZATION: International Institute for Applied Systems Analysis

AUTHOR: Keeney, Ralph L. and Raiffa, Howard

DATE: 1976

REPORT NUMBER/PUBLISHER: John Wiley and Sons, Inc.

ABSTRACT:

The emphasis of this definitive textbook on multi-attribute decision analysis is found in its subtitle: preferences and value tradeoffs. Once a decision problem has been formulated and modelled, two key features of the analysis arise--an uncertainty analysis and a preference (or value or utility) analysis. Whereas a great deal has been written on the uncertainty phase--statistical validation of the model, uses of historical and experimental data for inference, codification of judgmental estimates by the decision-maker or expert groups, etc.--little has been written about the value or preference side. The authors attempt to redress this imbalance. The relative desirability of alternatives or outcomes is expressed not by a single number, but by multiple numbers--some reflecting benefits, others reflecting costs. These are often expressed in incommensurable units. The authors say that the decision-maker cannot simply plug these incommensurate output performance measures into an objective function. They maintain that the decision-maker must think hard about the various value tradeoffs and about his attitude toward risky choices (the uncertainty phase); this leads them to suggest ways that the process can be systematically examined by dividing the overall, complicated choice problem into a series of simpler ones. The methodology relies on the principle of maximization of expected utilities, which are assigned to each attribute of the alternative outcomes and which may be averaged over a set of simulations run to reduce uncertainty. The book is best absorbed in small doses--and its use of many applications as illustrative examples suffers from the fact that it presumes considerable familiarity with that theory.

TITLE: Evaluating Management Information Systems

ORGANIZATION: N/A

AUTHOR: King, William and Rodriguez, Jaime

DATE: September 1978

REPORT NUMBER/PUBLISHER: MIS Quarterly

ABSTRACT:

The authors present a methodology for evaluating the value of management information systems and apply the method to a given application.

In short, the authors state that the value of any information system must be measured in four different categories: attitudes and value perceptions of the users, the degree of user behavior change brought about by the availability of the data, and whether the quality of decisions have improved. The article does not directly address the quality of data issue.

TITLE: Centralized Versus Decentralized Data Environment

ORGANIZATION: Veterans Administration

AUTHOR: Kuhn, Ingeborg

DATE: February 1986

REPORT NUMBER/PUBLISHER: U.S. Department of Commerce

ABSTRACT:

The author discusses the pros and cons of centralized versus decentralized data processing within an area of the Veteran Administration's Department of Medicine and Surgery.

The author states that this department of the VA is evolving to a decentralized environment and is successfully doing so with the aid of a good data dictionary and strong data administration policies.

The author summarizes the benefits of a decentralized environment as follows:

- Data is available for immediate local use.
- Greater incentive for accurate data capture (because the data are used on the local level).
- Capability to add unique local data needs.
- User defined data leads to increased validity.

The disadvantages of a decentralized environment are:

- Lack of central control over data element definition.
- Need for reconciliation between agency standards and user standards.

TITLE: Data Dictionaries Can Point the Way

ORGANIZATION: N/A

AUTHOR: Kull, David

DATE: July 1987

REPORT NUMBER/PUBLISHER: Computer & Communications Decisions

ABSTRACT:

The author discusses the advantages and current capabilities of several data dictionaries. In short, a data dictionary, if integrated with the software development process can save corporations hundreds of thousands of dollars per year in both development and operation and maintenance.

The author briefly discusses the need for a standard data dictionary and the different efforts under development. Most companies are waiting for IBM's lead.

TITLE: Information Resource(s) Management-IRM

ORGANIZATION: The KBL Group

AUTHOR: Levitan, Karen B.

DATE: 1982

REPORT NUMBER/PUBLISHER: Annual Review of Information Science and Technology

ABSTRACT:

Levitan reviews both the theory and practice of Information Resources Management (IRM); although, at the time, IRM was more of a direction and objective than a full-fledged practice. She discusses the various definitions and objectives of IRM, public-sector and private-sector activities, and associated methodologies. Among the techniques used for IRM are inventories (directories and data dictionaries); information flow analysis; information environment analysis (where the environment can be examined in terms of variables, including organization objectives, management communications, and types of decision-making processes); and accounting and budget. The various techniques used for cost accounting and budgeting of information resources include value/use approaches, which parallel the emphasis on user perceptions critical to the ISC task.

Levitan presents a thorough review of IRM, complete with names, organizations, and relevant documentation. She says that--while interest and activity are significant--there is as yet no IRM methodology, only suggested approaches. Our research suggests that there has been little change since the date of her paper.

TITLE: Databases and the Role of the Data Dictionary

ORGANIZATION: N/A

AUTHOR: Lobley, Don

DATE: March-April 1981

REPORT NUMBER/PUBLISHER: Cost and Management

ABSTRACT:

The author reviews the categories of data redundancy in MIS, the bad aspects of data redundancy, the purpose of a data dictionary, and the role of a database administrator.

Data redundancy obtains whenever the same data item is used to mean something different between organizations, or conversely, the same item has different names. The author states that the degree of this problem gets worse in organizations with several divisions, each of which developed their own ways of keeping data. The author points out that true redundancy is the same data duplicated within a system design to minimize response time in applications.

The author explains the importance of a database administrator as a "data gate-keeper" and describes the tools he uses to do his job. The DBA uses a data dictionary to maintain and control the use of data definitions within application systems. It contains descriptive information, technical characteristics and interrelationships between data items. The benefits of a good data dictionary are:

- Multiple updating processing is minimized
- Data values are consistent
- Cost and time are reduced in developing new applications and modifying old ones

TITLE: Cost Benefit Analysis of Intercomponent Data Access Alternatives

ORGANIZATION: Federal Computer Performance Evaluation and Simulation Center

AUTHOR: Ludington, E.; Perez, Christine; Targuini, Sally; and Yang, Arthur

DATE: February 1987

REPORT NUMBER/PUBLISHER: FEDSIM Report

ABSTRACT:

The authors present a cost benefit analysis of several alternatives to evaluate ways of sharing personnel data between several Army personnel components. The four alternatives studied were:

- Transmitting data manually, via tapes (the current mode of operations--the baseline),
- Accessing other component databases through remote terminals (remote access method),
- Accessing other component databases through a "transparent-heterogeneous access" mode whereby interfacing software is written to "link" databases resident on different machines and software, and
- Accessing other component databases through a "transparent-homogeneous" access mode whereby databases are "linked" to databases resident on the a hardware and software

The results of the analysis were in terms of cost/benefit ratios and payback periods which are shown below:

	<u>C/B RATIO</u>	<u>PAYBACK (YEARS)</u>
Remote Access	3.31	3
Transparent-Heterogeneous	1.26	6
Transparent-Homogeneous	2.24	4

Benefits and costs were assessed using the incremental improvement method. This involved calculating the incremental benefits and costs over the baseline approach, i.e., the manual tape transfer approach.

This study formed the basis of the current total Army personnel database. This database is a "mini" Army corporate database for Military Personnel Center (MILPERCEN), the Civilian Personnel Center (CIVPERCEN), the Army Reserve Personnel Center (ARPERCEN), and the National Guard. The Army selected the transparent-heterogeneous alternative.

TITLE: Decision-Making Among Multiple-Attribute Alternatives: A Survey and Consolidated Approach

ORGANIZATION: RAND Corporation

AUTHOR: MacCrimmon, K. R.

DATE: December 1968

REPORT NUMBER/PUBLISHER: RM-4823-ARPA

ABSTRACT:

In all types of decision situations, the alternatives from which a choice must be made are characterized by multiple attributes (or properties). For example, a weapon system may be characterized by performance, cost, availability date, etc. These characteristics may be further subdivided, e.g., performance may be considered relative to range, time, yield, vulnerability, and accuracy. As the number of relevant attributes and alternatives increases, the ability of the decision-maker to handle the problem decreases, and the information processing requirements may rapidly exceed the decision-maker's processing capacity.

This lengthy research paper provides a description and evaluation of several existing methods for helping decision-makers deal with the multiple-attribute decision problem. Methods that are reviewed include: Dominance, Satisficing, Maximin, Minimax, Lexicography, Additive Weighting, Effectiveness Index, Utility Theory, Tradeoffs, and Nonmetric Scaling. The author, after comparing the strengths and weaknesses of all these methods, recommends that an eclectic approach be adopted--which uses several of the methods most appropriate for the particular decision.

The paper is a good primer on the multiple attribute problem, which is somewhat relevant to the evaluation of MIS-based decisions--since the quality of these decisions needs to be assessed on the basis of multiple attributes.

TITLE: Evolution Towards Strategic Applications of Databases Through
Composite Information Systems

ORGANIZATION: Center for Information Systems Research

AUTHOR: Madnick, Stuart and Wang, Richard

DATE: May 1988

REPORT NUMBER/PUBLISHER: MIT, Sloan School of Management, CISR WP-170

ABSTRACT:

The authors present an evolutionary method of integrating disparate databases into a Composite Information System (CIS). A CIS is a system which integrates "independent" systems which may reside within and/or across organizational boundaries. In developing this strategy, the authors present four key principles which must be applied in order to achieve a successful CIS. First, the data must be separated from the processing. Second, the implementing software must contain special purpose languages to facilitate the construction of "integrated" applications. Third, the composite software must be able to invoke the processing component of the distant DBMS, be able to map between two databases with dissimilar types of data models, and convert data formats retrieved from the distant database into the processing format of the CIS. Fourth, the explicit recognition of the CIS environment.

TITLE: Integrating Disparate Databases for Composite Answers

ORGANIZATION: Center for Information Systems Research

AUTHOR: Madnick, Stuart and Wang, Richard

DATE: September 1987

REPORT NUMBER/PUBLISHER: MIT, Sloan School of Management CISR WP-171

ABSTRACT:

The authors discuss a number of the problems in attempting to integrate different databases, review some of the current technological solutions, and present a design to overcome these problems.

The authors state that the key idea of a CIS is integration for composite answers. Without integration, it is difficult, expensive, time-consuming, and error prone to obtain key information which may be distributed in databases located in different divisions of different organizations. The authors briefly review twelve different methods have been research to yield a CIS, and then present a case study which depicts the major data incompatibilities which cause these integration methods to fail.

Through the use of advanced DBMSs and Artificial Intelligence/Expert Systems, the authors present a theoretical technique of providing a CIS which allow "query by objectives". The authors present a vivid picture of the types of data problems that can occur with the example of three tour guide information systems which provide lodging amenity information. Each system collected "data" from the same source, but data was organized based upon different models used by the different designers. The authors present a theoretical approach, using artificial intelligence, to resolve data contradiction, inconsistency, and ambiguity as applied to integrating heterogeneous, physically distributed databases.

TITLE: The Data Dictionary: An Evaluation from the EDP Audit Perspective

ORGANIZATION: N/A

AUTHOR: Mannino, Michael; Solomon, Ira; and Vanecek, Michael

DATE: March 1983

REPORT NUMBER/PUBLISHER: MIS Quarterly

ABSTRACT:

The authors describe the features of current data dictionaries and evaluates potential benefits, and limitations of data dictionary systems from the perspective of the Electronic Data Processing (EDP) auditor.

TITLE: What is the Value of Investment in Information Systems?

ORGANIZATION: Information Systems, Land O'Lakes, Inc.

AUTHOR: Matlin, Gerald L.

DATE: September 1979

REPORT NUMBER/PUBLISHER: MIS Quarterly, Vol 3, No. 3

ABSTRACT:

If business decisions regarding expenditures for information systems are to be made rationally, then the title question must be posed and answered. The primary issue, in the process recommended in this paper, is placing dollar amounts on intangible benefits. Operational benefits such as reduction of staff have tangible dollar value to the company and are relatively easy to evaluate. The author maintains that difficulty in finding a value is not an appropriate reason for ignoring the quantification of intangible benefits; he proposes a valuation technique that involves close interaction and directed conversations with managers to derive approximate values. The approach is decidedly business-oriented and practical, not scientific or academic. Irrespective of the quality of the estimates, Matlin feels that the underlying objective of the process has to be an increased understanding of, and appreciation for, the benefits of information systems investment.

His technique of guided discussions includes preliminary reviews of corporate objectives to be met by the information systems and calculation of the percentage of objectives scheduled (and achieved). He then compares investment costs with values for objectives achieved to derive value/cost ratios. He describes four methods for use in the value computation: replacement cost, cost avoidance, related or analogous benefits, and comparison values. Finally, he demonstrates the use of the technique with an investment evaluation at Land O'Lakes, Inc.

TITLE: The Economics of Information and Optional Stopping Rules

ORGANIZATION: RAND Corporation

AUTHOR: McCall, John J.

DATE: October 1964

REPORT NUMBER/PUBLISHER: RAND P-2987

ABSTRACT:

This paper is a presentation of methods for measuring the economic value of information, where the problem is described as when to stop collecting information (or presumably, improve the quality of the information you have). The appropriate amount of information to purchase depends, of course, on the value of the information (in terms of added payoff) relative to its acquisition costs. Two concepts of statistical decision theory are reviewed: the expected value of perfect information and the expected value of sample information. The author lays out the mathematical framework for maximizing the decisionmaker's expected utility in two situational problems: replacement policy for an equipment which cannot be repaired (a repetitive decision); and capital/budgeting investments (a one-time problem). It is a complex framework, involving Markov processes and dynamic programming, with little real value to the Information Systems Command (ISC) task.

TITLE: Basis for Stamis Integration

ORGANIZATION: The MITRE Corporation

AUTHOR: Mendenhall, J. H.; Simko, M. J. and Stevens, W. B.

DATE: September 30, 1988

REPORT NUMBER/PUBLISHER: MITRE, WP-88W00388

ABSTRACT:

This internal working paper examines 30 Army Standard Management Information Systems (STAMIS) as far as their potential for reducing duplicate data, functions, and cost.

The authors examine the definition of integration as stated by DISC4. They explain the difference between functional, technical, programmatic, and program and resource integration. The authors address only functional and technical integration. In particular, they present selection criteria for functional and technical integration for the STAMIS. The authors also present lessons learned from the development of a proof-of-concept prototype integration effort which MITRE performed.

The authors present a clear analysis of the factors which should drive integration of Army STAMIS and illustrate the nature of problems encountered when attempting to do so.

TITLE: Cost Justification Report Guide

ORGANIZATION: ADPAC Corporation

AUTHOR: N/A

DATE: 1989

REPORT NUMBER/PUBLISHER: N/A

ABSTRACT:

This guide is a sample report for structuring cost-benefit justification of data administration software, more specifically, ADPAC's software productivity tools, PM/SS. Although couched in terms of a specific product's projected performance, it contains a descriptive listing of functions and potential cost savings for each--data administration, development, quality assurance, and maintenance. The cost benefit analysis compares the cost reductions (i.e., "benefits") achieved through automated data standardization, data dictionary preparation, impact analysis, etc. with the actual costs of the software to derive a return on investment. Sample calculations and likely percentage savings are contained in a separate appendix and provide some general, order-of-magnitude estimate of potential cost savings by category.

TITLE: AIRMICS Research Status Report

ORGANIZATION: U.S. Army Institute for Research in Management Information,
Communications, and Computer Sciences (AIRMICS)

AUTHOR: N/A

DATE: September 1988

REPORT NUMBER/PUBLISHER: AIRMICS

ABSTRACT:

This Research Status Report summarizes ongoing research tasks and technical support activities within DOD that are being performed by the AIRMICS, the research arm of the U.S. Army Information Systems Command (ISC). AIRMICS provides direct support to the Information Systems Engineering Command (ISEC) technical staff, ISC Headquarters and its subordinate commands, and to the Program Executive Officers (PEOs) and Program Managers (PMs) in the Information Mission Area (IMA). The report covers research activities in three functional areas:

- Computer and Information Systems
- Communications and Network Systems
- Advanced Concepts and Technology Integration

Research is underway in several areas related to the Data Economics task-decision support, information management, and distributed systems. The specific scope of these studies were not, however, relevant to the task.

TITLE: Weather Information and Economic Decisions: A Preliminary Report

ORGANIZATION: RAND Corporation

AUTHOR: Nelson, Richard R. and Winter, Sidney G., Jr.

DATE: August 1960

REPORT NUMBER/PUBLISHER: RM-2620-NASA

ABSTRACT:

This study examines ways in which weather information enters into economic decisions. It develops methods for determining the dollar value of information varying in quality from climatological probability to the "perfect" forecast; and it illustrates how these methods can be applied to a number of typical (weather-sensitive) business decisions. The framework is applied to a roofing-construction problem where good forecasts enable the decisionmaker both to avoid the losses of doing the job under unfavorable conditions and to exploit favorable conditions when they occur.

The authors set up a matrix of actions and possible weather states (with probability distributions based on historical and forecast data) within which the decision-maker attempts to maximize economic payoff. Then, an improvement in forecasting information is achieved (e.g., through weather satellites), which reduces uncertainty, and thereby increases decisionmaker payoff.

The study is an interesting application of expected value analysis which directly attributes value to information.

TITLE: Guidelines for an Economic Analysis for the Army Corporate Database

ORGANIZATION: The MITRE Corporation

AUTHOR: Pueschel, Larry

DATE: February 1988

REPORT NUMBER/PUBLISHER: MITRE

ABSTRACT:

This draft MITRE Technical Report discusses the concept of a cost benefit analysis and presents a five step procedure for conducting one. The analysis is written as a guideline for conducting a cost benefit analysis for the Army Corporate Database concept.

Here is a summary of the cost benefit analysis tasks and subtasks:

1. Define the problem

- a) Define scope/baseline problems
- b) Identify objectives
- c) Identify alternatives and solutions
- d) Identify assumptions and definitions
- e) Discuss format and results

2. Design the analysis

- a) Determine decision form from problem structure
- b) Identify benefits and measurement
 - i) Determine list of objectives
 - ii) Determine resources impacted if objectives are met
 - iii) Develop resource impact matrix--by alternative
 - iv) Develop benefit transformations
 - v) Validate benefits
- c) Identify costs and measurement methodology
- d) Identify sensitivity analysis methodology
- e) Identify data to be collected

3. Collect data

4. Analyze data

5. Present results

The author also summarizes some of the important hypotheses made about the current stove-pipe data environment. These problem statements were obtained from the Army Corporate Database Implementation Strategy Plan and the Corporate Database Concept and Strategy papers. Some of these problems are presented below:

1. There is significant inappropriate use of data from stove-pipe sources.
2. Lack of consistent, accurate data contributes significantly to the improper allocation of resources.
3. Lack of consistent, accurate data significantly impairs the Army's decision making ability, contributing to a loss in readiness.
4. There is a high degree of unnecessary data redundancy within Army systems.
5. There is a significant loss in horizontal communication, resulting in a lack of access to key information.
6. There is a loss of system impact observability and controllability of actual events due to self-limiting definitions.
7. Army data are not synchronized. Aggregate data are often disparate, inaccurate, and incapable of synchronization.
8. Army data are not standardized. Lack of standardization contributes significantly to the inability of Army personnel to properly perform their jobs.
9. Army data are not timely.

The author also presents a list of preliminary "benefits" which would accrue from a shared ACDB environment. Most of these benefits are:

1. Increased access--users should be able to access all authorized information services from essentially any terminal.
2. "Go to War" capability--users will be able to use common systems, the same functional applications, and the same set of automation skills in all stages of conflict (peacetime, crisis, transition to war, and wartime).

3. From the users' point of view, the network will be transparent, within unavoidable limits.
4. Database will have standard data elements.
5. Database will be synchronized in the representation of data; e.g., time period represented or actual versus projected date identified.
6. Increased access causes additional multi-level security.
7. Maximized system survivability.
8. Functional integration--among tactical, strategic, and sustainment operations.
9. Increased timeliness of data.
10. Increased accuracy of data.
11. Flexibility in spite of organizational changes.
12. Enhanced mission effectiveness.
13. Better resource allocation.
14. Improved programming efficiency.

TITLE: Putting Top-Down and Bottom-Up Analysis Together

ORGANIZATION: Merrill Lynch

AUTHOR: Rice, Daniel and Laufer, Sandra

DATE: December 1988

REPORT NUMBER/PUBLISHER: Database Programming and Design

ABSTRACT:

The authors discuss the idea of resolving existing "bad data" on systems using a current automated tool. The authors describe some of the useful logical data modeling tools in use for new applications, but point out that few do little to resolve existing data problems. Many firms avoid the resolution of existing data problems because it is too hard due to sketchy documentation usually out of date and back tracking data elements, their homonyms, synonyms, and data dependencies out of source code. In spite of the arduous task of data resolution, the authors argue that the payoff can be very substantial. The authors guess that nearly 50 percent of the software maintenance effort could be eliminated if the data problems were fixed. Furthermore, the authors state data redundancy in many systems is as high as 75 percent which results in much more disk space than needed.

The authors recommend fixing data problems in existing systems through the use of automated tools. They state the use of CASE tools can help the problems from recurring in the first place, but the CASE technology needs a "reverse engineering" capability.

TITLE: What If Utility Functions Do Not Exist?

ORGANIZATION: RAND Corporation

AUTHOR: Roberts, Fred S.

DATE: September 1970

REPORT NUMBER/PUBLISHER: R-528-ARPA

ABSTRACT:

This paper discusses alternatives to the strict decision-making goal of ranking all alternatives or, equivalently, obtaining an utility function. For many military problems, the best information available is the combined judgments of experts. Often, however, the preferences of decision-makers are too inconsistent or ambiguous to permit a complete ranking of alternatives or an utility function. Optional methods for dealing with such a situation include: (1) finding a procedure through which preferences can be modified to obtain a utility function (e.g., Delphi); (2) using the utility assignment that best approximates a utility function (e.g., averaging rankings on different facets of the decision); or (3) modifying the demands on utility functions. This study emphasizes the third method, and describes it in terms of techniques from the theory of measurement. It is a modification, in the usual utility sense, because it tries to derive through mathematical logic a probabilistic consistency to decision-making-not absolute consistency or rationality. The paper is useful in that it presents alternate methods which can be used to determine relative preferences in decision outcomes; however, it does not furnish any insights in terms of actual evaluation of these outcomes or in terms of the efforts of information on outcomes.

TITLE: Data Dictionaries and Data Administration

ORGANIZATION: N/A

AUTHOR: Ross, Ronald G.

DATE: 1981

REPORT NUMBER/PUBLISHER: American Management Association

ABSTRACT:

This is a text book dealing with the subject of data resource management, the data administrator and his principal administrative tool, and the data dictionary.

The book is divided into four parts. Part I deals with explaining how the proper use of a data dictionary can make or break a data resource development project. Part II is a guide to data dictionary acquisition in terms of major considerations. Part III deals with the management of data administration and Part IV discusses the inter-relationship of the data dictionary to the entity's DataBase Management System (DBMS). The author explores the capabilities of the DBMSs on the market.

The book contains useful definitions in the data resource area. It also addresses the problems associated with a distributed data processing environment, particularly as it pertains to the Data Administrator's responsibilities.

TITLE: Alternative Measures of System Effectiveness: Associations and Implications

ORGANIZATION: Indiana University

AUTHOR: Srinivasan, Ananth

DATE: September 1985

REPORT NUMBER/PUBLISHER: MIS Quarterly

ABSTRACT:

The objective of this article is to examine, in more detail, the relationship between user perceived effectiveness measures (i.e., user satisfaction) and behavioral measures of MIS effectiveness (i.e., system use) to test the implication that the two types of measures are positively correlated (and thereby substitutable). Twenty-nine firms were surveyed to collect data on the uses of both measures and an associative analysis was compiled on the following dimensions of the measures:

● Perceived Measures
(USERS)

- report content
- report form
- problem solving
- input procedures
- system stability

● Behavioral Measures
(MIS System)

- frequency of use
- time per session
- number of reports
- user type

Performance measures for each of these dimensions are presented, and pairwise correlation indices are derived on the basis of survey results. The author concludes that, in general, the correlation between user satisfaction and system effectiveness is not consistent (sometimes positive, sometimes negative). This contradicts most MIS literature which claims that the two perspectives track together. Besides the interesting conclusion, this paper is important for its exposition on user measures of value or effectiveness.

TITLE: Value-Added Processes in Information Systems

ORGANIZATION: Syracuse University

AUTHOR: Taylor, Robert S.

DATE: 1986

REPORT NUMBER/PUBLISHER: Ablex Publishing Corporation

ABSTRACT:

This textbook addresses a persistent theme in information requirements analysis--development of a pragmatic measure of information utility or value. A major argument of the author is that the value of the information is user based, i.e., the value of information cannot be measured in precise terms prior to its use. Value is therefore conveyed by the user, not the system designer. The value-added approach relies on an evaluation of the information beneficiaries and their need for and mode of use of information. Value-added activities in information systems are those processes that produce, enhance, or otherwise strengthen the potential utility of information messages in the system. Several of the values that the author highlights fall under the category of quality, which includes data accuracy, comprehensiveness, currency, reliability, and validity. Value added in the context of data quality means that the user obtains a residual value over and above the cost of implementing improvements in data quality; this is roughly analogous to the hypothesis of the data economics task that states that higher quality of MIS data leads to better (however valued) decisions.

The author presents a value-added model that sketches out a structure for the collection and organization of data on the user environments, system design, system outputs, cost-benefit data, and outcomes/effects of information on performance. This is a useful tool for conducting an impact evaluation of information systems. Moreover, there are particularly relevant sections of the book which specifically look at information in the decision context (programmable and nonprogrammable decisions) and at information versus productivity considerations (costs and benefits). Although these are descriptive rather than quantitative in nature, these chapters are very helpful in formulating a model of the evaluation process.

In summary, the book is an excellent introduction and essential reference for any researcher in the information value field.

TITLE: Army Data Element Standardization Program

ORGANIZATION: USA Audit Agency

AUTHOR: USA Audit Agency

DATE: March 28, 1986

REPORT NUMBER/PUBLISHER: USA Audit Agency

ABSTRACT:

This report evaluates whether the Army's Data Element Standardization Program has been adequately implemented. The report states that the program was not adequately implemented, that Army policies and procedures resulted in fragmented resource management, proliferation of systems which did not use standard data elements, and Army policies that emphasized functional control instead of data management.

The report also found that the Army's catalog of Standard Elements and codes did not contain information required to control or manage data. Further, the report states that functional proponents and system developers were not using the standard catalog, nor were they proposing new data elements.

The report makes several recommendations which undoubtedly led to the development of AR-25-1 and AR-25-9 which replaced the catalog of Standard Elements (AR-18-xx). The report contains a concise history of the Army data standardization efforts.

TITLE: AR 25-9: Army Data Management and Standards Program

ORGANIZATION: USA, DISC4

AUTHOR: USA, DISC4

DATE: July 8, 1988

REPORT NUMBER/PUBLISHER: USA, DISC4

ABSTRACT:

This regulation implements the management of data for the Information Mission Area and the Army Information Resources Management Program (AR 25-1). It prescribes the policies, responsibilities and concepts of operation for the management of data used in manual and automated information systems throughout the Army. The purpose of the regulation is to provide the framework that will transition the Army from individual and incompatible systems to the objective, shared database environment.

The regulation contains two chapters and several appendices. Chapter 1 describes the purpose, concepts, goals, objectives of the regulation and assigns responsibilities to Army components. Chapter 2 contains information for standardizing Army data elements. The appendices provide guidance on the content of entries for the Army data encyclopedia.

TITLE: The Army Data Architecture

ORGANIZATION: USA, DISC4

AUTHOR: USA, DISC4

PUBLICATION/PUBLISHER: USA, DISC4

DATE: 7 May 1986

ABSTRACT:

This document contains the Army's data management framework, or guiding rules for organizing data to support the way an organization does or wants to do business. The purpose of the architecture is to facilitate the identification of data redundancies and inconsistencies and move the Army to a data sharing environment.

The Data Architecture is an entity-relationship model that identifies data by segregating it into entities: those persons, places, things, concepts, events or activities about which an organization wants to keep data. This Data Architecture depicts the fundamental data relationships among those entities. The Data Architecture provides the basis for design of subject area databases.

TITLE: Management of Federal Information Resources

ORGANIZATION: Office of Management and Budget

AUTHOR: U.S. Government

REPORT NUMBER: Circular No. A-130

DATE: December 1985

ABSTRACT:

This circular establishes policy for the management of Federal information resources, based on the broad mandate established in the Paperwork Reduction Act of 1980. Procedural and analytic guidelines for implementing specific aspects of these policies--e.g., records on individuals, security, and cost accounting--are attached as appendices. The final appendix is an analysis of key sections of the circular No. A-130; this is the most useful section of the circular in terms of setting down more specific requirements. The document overall has little value to the ISC task, however, beyond describing the OMB oversight responsibilities and the policies to be followed by Federal agencies in improving the cost-effectiveness of their information management activities.

TITLE: Paperwork Reduction Act of 1980

ORGANIZATION: U.S. Congress

AUTHOR: U.S. Government

REPORT NUMBER: Public Law 96-511

DATE: 11 December 1980

ABSTRACT:

The purpose of this act is to minimize paperwork and cost to the Federal Government of collecting, maintaining, using, and disseminating information; to coordinate, integrate, and maximize usefulness of information collected by the government; and to ensure that ADP and telecommunications technologies are acquired and used in a cost-effective manner. The act's provisions are administered by the OMB, who is enjoined to achieve specific levels of information burden reductions and prepare plans to that effect. Each Federal agency is assigned responsibilities to support OMB through an inventory of its information systems and to periodically review its activities and procedures for managing its information. The act was instrumental in identifying information as a "resource" to be managed - including cost and value. This spawned the Information Resources Management (IRM) concept, which has resulted in a number of interesting information analysis applications.

TITLE: Preparation and Submission of Budget Estimates

ORGANIZATION: Office of Management and Budget (OMB)

AUTHOR: U.S. Government

REPORT NUMBER: Circular No. A-11

DATE: June 1988

ABSTRACT:

This document provides detailed instructions and guidance on the preparation and submission of annual budgets and associated materials. It is the most recent version, updated for the Fiscal Year 1990 (FY90) budget process. Of special interest is Section 43.1, "Data on Acquisition, Operations, and Use of Information Technology Systems," which describes the reports required by OMB to fulfill its oversight role in information technology as prescribed in the Paperwork Reduction Act of 1980. These reports, which are required from all agencies that obligate more than \$100K in PY, CY, or BY for Information Technology activities are:

- Obligation report
- Narrative statement and agency acquisition plan, and
- Benefit-Cost analysis (to be submitted with the initial budget submission).

The Benefit-Cost analysis requirements are relevant to the ISC task, since it describes, in some detail, the guidelines for these analyses. The analyses are required for information technology initiatives meeting any of the following:

- Total cost exceeds \$5M; or cost in any one year exceeds \$2M
- Proposed acquisition is a significant initiative that will install, automate, or modify a major information system of the agency.
- Any system at the discretion of OMB.

Other guidelines or requirements include a demonstrated 10 percent return on investment (or positive present value, using a 10 percent discount rate). Initiatives not meeting this standard shall include substantial additional justification for funding, based on a specific statutory requirement or quantifiable, but non-economic improvements to the agency's ability to perform its mission.

TITLE: Real World Data Management

ORGANIZATION: N/A

AUTHOR: Van Rensselaer, Cort

DATE: October 1988

REPORT NUMBER/PUBLISHER: Computer Decisions

ABSTRACT:

The author of this article is a 40-year veteran of the Hewlett-Packard Corporation who offers several methods to maintain data integrity.

The author claims that, although technology is currently available to help organizations solve data management, they generally do not know how to use it. Even if companies know how to use the tools (i.e., data dictionaries, database management systems, data entity relationships modeling, and CASE), most will fail because of the lack of sound business fundamentals. The author describes the lack of sound business fundamentals which have worked for HP.

In describing these policies, the author offers some facts on why strictly technically-oriented solutions to the "bad data" problem have not worked. The author cites a study performed by MIT's Center of Information Research that confirms the premise that most successful data management efforts are driven by fundamental considerations directly related to business needs. Five of the 20 companies undertook strategic data planning projects for a completely "shared" data environment and each failed. These companies were unsuccessful because the data had different meanings to different entities, and the efforts took too long--by the time some understanding was believed, the business as a whole had changed and, hence, the data was rendered meaningless.

The policy recommendations to fix the "bad data" problem are:

- Set up a distributed environment, but make the lowest operating using entity responsible and accountable for "their data"
- Limit the scope of a data sharing approach
- Manage cross-functional data elements centrally
- Use industry data definitions standards whenever possible

- Use reference databases as sources of data
- Make decision data a by-product of transaction systems

TITLE: Information Economics

ORGANIZATION: Institute for Software Engineering

AUTHOR: Vincent, D. R.

DATE: 1983

REPORT NUMBER/PUBLISHER: Information Executive, Vol.1

ABSTRACT:

This paper introduces the notion of information economics, a discipline concerned with the production, distribution, and consumption of information and knowledge. Vincent contends that data processing has evolved from a service, overhead function to become--as information--a corporate asset. Since it is not unusual to see organizations dedicating 25 percent of their resources to the creation and maintenance of information, he foresees that corporations will soon include information in their financial reports--replete with amounts spent on information investment, return on investment, and productivity gains expected in the current and future reporting periods. There is a very good discussion of the pertinent investment costs within each of these information coordination and management activities:

- Storage of the information asset
- Management of the information database
- Information processing
- Information transfer

Although he states that a comprehensive understanding of the value of information will result in increasing gains in corporate productivity, his focus in the article remains on costs, not benefits or value.

TITLE: Cost Justifying the Data Administration Function

ORGANIZATION: Bank of America

AUTHOR: Voell, Ronald F.

DATE: February 1986

REPORT NUMBER/PUBLISHER: National Bureau of Standards (NBIR-3324)

ABSTRACT:

This paper provides guidance to data administrators who are asked to estimate the cost-effectiveness impact of the data administration function on management performance. Some benefits may be difficult to quantify (e.g., improved data accuracy), but others, such as reduction of storage space because data redundancy is decreased, should be quantifiable and reflected in reduced costs. The key to Voell's commentary is that organizations should examine each benefit in detail to see if it is really intangible, and then see if there is not data within or outside of the organization that can provide some quantifiable insights.

Voell discusses eight possible methods for measuring the cost benefits of data administration. These range from simply attributing (not quantifying) an effect (more efficient labor use, less data redundancy) to labor cost analysis and data modeling. Voell reports on the results that were measured in actual business settings and contrasts these findings with those of other researchers. The emphasis of the paper is not these numbers (which range from five to 15 percent savings for development phases, 80 to 85 percent for operations), but on techniques that can be adapted and used in all organizations to quantify the actual benefits being achieved by data administration.

Since data administration encompasses many of the data quality/standardization issues of interest to the ISC, this overview provides some interesting methodological options for at least capturing the "efficiency-based" benefits. User effectiveness not directly addressed.

APPENDIX B

INTERVIEW SUMMARIES

Over 45 organizations and individuals were contacted and interviewed to discuss their experience and insights into the issues of data redundancy, information system evaluation techniques, and impact of data/information quality on decision-making. A brief summary of 32 of these discussions are included in this appendix, arranged in alphabetical order of the person interviewed.

AGENCY/ORGANIZATION OF CONTACT: Stanford University

INDIVIDUAL CONTACTED: Dr. Kenneth Arrow
Head of the Center for Research in Economic
Efficiency
(415) 723-9165

DATE: 18 January 1989

DISCUSSION:

1. The MITRE study group called Dr. Kenneth Arrow upon the recommendation of Mr. Randy Simpson. Dr. Arrow is a mathematical economist and a Nobel prize winner in Economics.
2. Dr. Arrow has not had any personal involvement with work that ties data quality into a decision-making framework. However, he is aware of work that has been done in this area at the Massachusetts Institute of Technology (MIT).
3. Dr. Arrow's area of interest is the decision-making process and how decision-makers use information to form conclusions and optimize results based on theories, rational behavior, and revealed interpersonal preferences.

AGENCY/ORGANIZATION OF CONTACT: General Services Administration
(GSA)/Federal System Integration and
Management Center (FSIMC)

INDIVIDUAL CONTACTED: Ms. Carol Campbell
Operations Research Analyst
(703) 756-4120

DATE: 9 January 1989

DISCUSSION:

1. The MITRE study group called Ms. Carol Campbell upon the recommendation of Mr. Spiros Coutavas of the MITRE Corporation. Ms. Campbell is an Operations Research Analyst with the General Services Administration (GSA)/Federal System Integration and Management Center (FSIMC).
2. Ms. Campbell is not aware of any information in the study area.

AGENCY/ORGANIZATION OF CONTACT: Air Force/Directorate of Architecture and Technology, Information Architecture Branch (AF/SCTIA)

INDIVIDUAL CONTACTED: LtCol W. Cato
(202) 695-9934

DATE: 27 January 1989

DISCUSSION:

1. The MITRE study group called LtCol W. Cato looking for a point of contact in the Air Force/Directorate of Architecture and Technology, Information Architecture Branch (AF/SCTIA).
2. LtCol Cato is not aware of any studies or reports directed at assessing impact of poor and inconsistent data on decision-making. However, within his organization there are efforts to develop a standardized data dictionary.
3. LtCol Cato recommends that the study group contact Mr. B. Nguyen within his organization.

AGENCY/ORGANIZATION OF CONTACT: Syracuse University

INDIVIDUAL CONTACTED: Ms. Sharon Caudle
Information Resource Management (IRM) Specialist
(315) 443-5608

DATE: 18 January 1989

DISCUSSION:

1. The MITRE study group called Ms. Sharon Caudle upon the recommendation of Mr. Morey Chick of the General Accounting Office (GAO). Ms. Caudle is an Information Resource Management (IRM) specialist on the faculty at Syracuse University. Ms. Caudle has done work on the value of information principles and has looked at the use of IRM in the Federal government for the National Association of Public Administrators (NAPA).
2. Ms. Caudle recommended that the study group review the following references:
 1. Value Added Processes and Information Systems by Robert Taylor.
 2. Infomap by Cornelius Burke and F. Horton.
 3. Various issues of the "Information Management Review".
 4. Various issues of the "Information Resources Management Journal".

AGENCY/ORGANIZATION OF CONTACT: General Accounting Office

INDIVIDUAL CONTACTED: Mr. Morey Chick
Assistant Director
(202) 275-0472

DATE: 18 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Morey Chick after reading an article he wrote. Mr. Chick works in the Information Management Office of the General Accounting Office (GAO). He has written a number of articles on the value of information to management, especially in the federal government.
2. Mr. Chick defined information as what comes out, data as what goes in, and metadata as information about data. The entire field is referred to as Information Research Management (IRM) and the objective is to maximize the usefulness of information.
3. Mr. Chick has examined and written about the impact of information problems such as data quality and duplication. He has traced the type of problems that have appeared in various government agencies. He feels his framework should apply to cleaning up the data that became the information used by decision-makers in the Army.
4. Mr. Chick feels that the Paperwork Reduction Act sets policy and objectives that are key to information improvement. Requirements are presented for cost-benefit studies for new systems.
5. Mr. Chick will send the study group a copy of a human safety article that examines why systems with information problems fail.
6. Mr. Chick recommends that the study group contact Ms. Helen Ebenfield at the National Science Foundation (NSF) and Ms. Sharon L. Caudle at Syracuse University.

AGENCY/ORGANIZATION OF CONTACT: Evaluation Technology Inc.

INDIVIDUAL CONTACTED: Mr. Russell Coile
Analyst
(408) 372-3439

DATE: 9 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Russell Coile upon the recommendation of Mr. Spiros Coutavas of the MITRE Corporation. Mr. Coile is an analyst with Evaluation Technology Inc.
2. Mr. Coile recommended that the study group use Dialog to query: the topic of errors in databases, how to detect errors, how are errors introduced into databases, what are error rates, the probability of errors in the transmission of data, and the topic of data reliability.

AGENCY/ORGANIZATION OF CONTACT: Data Administration, Inc.

INDIVIDUAL CONTACTED: Steve Cooper
President, DAI
(703) 684-4422

DATE: 18 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Cooper at the suggestion of Mr. Bruce Haberkamp of the Office, Director of Information Systems for Command, Control, Communications & Computers (DISC4). Mr. Cooper is involved in developing and producing software which will measure the degree of data redundancy within information systems.
2. Steve explained his firm is involved in what the Navy calls "data scrubbing". His firm had done some work for the Fleet Materiel Support Office in Mechanicsburg, Pa. In particular, his firm used two software programs to do this. One is called "Data Mapper" and the other "Data Expediter" which are Data Administration support tools. They were written primarily for Database applications with COBOL application programs. Data Expediter looks for a key word in context and key words out of context and then measures the extent of data redundancy.
3. The firm has manually performed impact analysis of data redundancy as well as estimating the cost of "bad data". By impact analysis he means they costed out the amount of "waste cost" by figuring out how much extra time a maintenance programmer spends due to data redundancy. The firm is also attempting to develop Cost Estimating Relationships (CERs) which would tell managers, based on certain data characteristics of the system, the extent of data redundancy. The cost of bad data they estimated was specific to a Navy supply application.
4. Mr. Cooper said he would send us some literature on their software product and services.

AGENCY/ORGANIZATION OF CONTACT: The MITRE Corporation

INDIVIDUAL CONTACTED: Ms. Twyla Courtot
Member of the Technical Staff
(703) 883-7343

DATE: 20 January 1989

DISCUSSION:

1. The MITRE study group called Ms. Twyla Courtot on the recommendation of Mr. David Stowell of The MITRE Corporation and Mr. Steven Cooper of DAI. Ms. Courtot is with the MITRE Washington Software Center.
2. Ms. Courtot is not aware of any studies that have examined the quantification of the cost of redundant or inconsistent data within Management Information Systems (MISs) and the related costs of bad decisions.
3. Ms. Courtot thought that the study group should be able to locate numerous articles addressing the additional processing time and maintenance costs resulting from the confusion of data elements within a system. The additional processing time is a result of longer access times.
4. Ms. Courtot recommended that the study group attend the Computer Aided Software Engineering (CASE) Symposium held at The MITRE Corporation on 23 January 1989 and speak with Mr. Peter Johnson of ASYST Technologies. ASYST is marketing an automated data dictionary. Ms. Courtot thought Mr. Johnson might have some experience with the study topic in the commercial sector.

AGENCY/ORGANIZATION OF CONTACT: Data Standards, Information Systems
Engineering Command (ISEC)

INDIVIDUAL CONTACTED: LtCol Robert DiBona
Chief, Policy Plans and Integration Division
(703) 664-3986

DATE: 17 January 1989

DISCUSSION:

1. The MITRE study group called LtCol Robert DiBona on the recommendation of Mr. James Glymph of Data Management Directorate (DMD), Information Systems Engineering Command (ISEC). LtCol DiBona is in charge of the ISEC Standard Army Management Information System (STAMIS) Modernization (STAMOD) program.
2. LtCol DiBona did not know of any studies which may have addressed the quantitative aspects of bad decisions which result from redundant or inconsistent data.
3. LtCol DiBona did not know of any studies which addressed the cost effectiveness of the Army Corporate Database (ACDB) concept. He did say that the reason the ACDB concept was abandoned was that it would take too long to implement and was just too big. The Army is currently moving toward a shared database environment. All new systems will be measured by ten criteria as designated in Army Regulation (AR)-25-9, Army Data Management and Standards Program. ISEC has the responsibility to evaluate these new systems, determine whether or not they comply, and stipulate compliance as necessary.
4. LtCol DiBona recommended that the study group contact Maj. C. A. Harris, Office of the Director of Information Systems for Command, Control, Communications, and Computers (DISC4) with regard to a DOCMOD study.
5. LtCol DiBona recommended that the study group contact Col Gerard Thrash, United States Army Force Development Agency, who might know of some studies which may have been done to justify the Force Builder project.

AGENCY/ORGANIZATION OF CONTACT: Management Information and Systems
Department, University of Texas

INDIVIDUAL CONTACTED: Dr. James Dyer
Professor
(512) 471-5278

DATE: 17 January 1989

DISCUSSION:

1. The MITRE study group called Dr. James Dyer upon the recommendation of Mr. Randy Simpson. Dr. Dyer is in the Management Information and Systems Department of the University of Texas. Dr. Dyer has experience in the fields of risk analysis and the diagnostic/medical consequences of risk and uncertainty.
2. Dr. Dyer agreed to send the study group reference material.
3. Dr. Dyer recommended that the study group contact Dr. Andrew Winston of the University of Texas.

AGENCY/ORGANIZATION OF CONTACT: National Science Foundation (NSF)

INDIVIDUAL CONTACTED: Ms. Helene Ebenfield
Economist
(202) 357-7601

DATE: 25 January 1989

DISCUSSION:

1. The MITRE study group called Ms. Helene Ebenfield upon the recommendation of Mr. Morey Chick at the General Accounting Office (GAO). Ms. Ebenfield is an Information Resource Management (IRM) specialist for the National Science Foundation (NSF).
2. Ms. Ebenfield does not feel that the study group will find quantification in the areas of the value of information or its impact on decisions. However, there are many descriptive papers that discuss these issues. Little work has been done on the benefit side of information.
3. Ms. Ebenfield recommended that the study group contact Dr. Daniel Newlon in the Economics Section of the NSF. Dr. Newlon is involved in econometric modelling and the measurement of error in data. Dr. Newlon recommended that the study group contact Mr. Larry Rosenberg in the Information Technology and Organization Section of NSF.

AGENCY/ORGANIZATION OF CONTACT: Army Institute for Research in Management
Information, Communications, and Computer
Sciences (AIRMICS)

INDIVIDUAL CONTACTED: Mr. Michael Evans
Program Manager
(404) 894-3104

DATE: 19 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Michael Evans upon the recommendation of LtCol William Meyers of the United States Army Information Systems Command (ISC). Mr. Evans works at the Army Institute for Research in Management Information, Communications, and Computer Sciences (AIRMICS), a research and development arm of ISC.
2. Mr. Evans will provide the study group with project sheets documenting their ongoing research.

AGENCY/ORGANIZATION OF CONTACT: Office for Naval Research

INDIVIDUAL CONTACTED: Mr. Neil Glassman
Mathematician
(703) 696-4313

DATE: 10 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Neil Glassman upon the recommendation of Dr. Thomas Varley of the Institute for Defense Analyses (IDA). Mr. Glassman is a mathematics oriented decision analyst with the Office of Naval Research (ONR).
2. Mr Glassman feels that the general topic of this study is "impossible".
3. Mr. Glassman recommends contacting the following people:
 - a. Mr. Kenneth Arrow at Stanford University.
 - b. Mr. Randy Simpson, formerly at ONR and presently a private consultant.
 - c. Mr. Andre Von Tilburg at ONR.

AGENCY/ORGANIZATION OF CONTACT: Data Management Directorate (DMD),
Information Systems Engineering Command
(ISEC)

INDIVIDUAL CONTACTED: Mr. James Glymph
Director
(703) 664-7134

Ms. Mary Jo Matera
Deputy Director
(703) 664-3224

DATE: 27 December 1988

DISCUSSION:

1. The MITRE study group met with Mr. James Glymph and Ms. Mary Jo Matera upon the recommendation of LtCol William Reyers of Information Systems Command (ISC).
2. Mr. Glymph does not feel that data standardization is essential or cost effective. As long as data are cross-walked as aliases, it is not worth the time or money to revise 15 databases to give elements the same name. Further, Mr. Glymph does not feel that there is a need for an Army Corporate Database (ACDB) and that one is not technologically possible.
3. The Data Management Directorate (DMD) is the database administrator for the Army. They are responsible for implementing Army Regulation (AR)-25-9, Army Data Management and Standards Program for ISC. There are 18 information class proponents that clear data elements for use in the various databases. DMD reviews and has approval over changes or inclusions of new data elements in Army databases.
4. DMD is not performing any cost/benefit or other studies that will justify the ACDB concept. DMD is developing Data Element Dictionaries and a Department of the Army Pamphlet that will implement AR-25-9, Army Data Management and Standards Program.
5. The United States Army Institute for Research in Management Information Communications and Computer Sciences (USAIRMICS) does academic research for ISEC and is based out of Georgia Institute of Technology.
6. Mr. Glymph has written a paper documenting the process 10 to 15 mid-level government employees (GS-15s and O-6s) go through to make decisions.

7. Mr. Glymph and Ms. Matera recommended that the study group contact the following people:
 - a. Mr. Greg Hoffman at the Office of the Assistant Secretary of the Army (Financial Management) (ASA(FM)) with regard to the extent of data redundancy within Army databases.
 - b. Mr. Jerry Cooper or LtCol Robert DiBona, Data Standards, Information Systems Engineering Command (ISEC), regarding the Standard Army Management Information System (STAMIS) Modernization (STAMOD) program.
 - c. Mr. Fred Simmons of the Database Standards Office, regarding an Army Audit Agency report on the problems associated with bad data.
 - d. Dr. Edgar Sibley of George Mason University. Dr. Sibley may know of work that has been done or other points-of-contact in the area of data redundancy.
 - e. Col Winkler, Director of Information Management, Personnel Information Systems Command (PERSINCOM). Mr. Glymph feels that maintaining "good" data is a problem with regard to updating data. Col Winkler can address the impact bad data has on systems.
8. Mr. Glymph and Ms. Matera recommended that the study group review the Management Information System (MIS) Quarterly and the Database Newsletter for papers that address the study issues.

AGENCY/ORGANIZATION OF CONTACT: Office of the Director of Information
Systems for Command, Control,
Communications, and Computers (DISC4)

INDIVIDUAL CONTACTED. Mr. Bruce Haberkamp
Analyst
(202) 694-0754

DATE: 18 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Bruce Haberkamp on the recommendation of LtCol Robert DiBona of Information Systems Engineering Command (ISEC) and Col Gerald Thrash of the United States Army Force Development Agency. LtCol DiBona and Col Thrash thought Mr. Haberkamp would have knowledge of studies that addressed the costs of redundant and inconsistent data. Further, they thought his office might have statistics documenting the Army Corporate Database (ACDB) concept.
2. Mr. Haberkamp acknowledged that Maj C. A. Harris and Maj Holt, both on his staff, were the appropriate individuals to speak with. Maj Harris had been identified by LtCol DiBona as having a DOCMOD study.
3. Mr. Haberkamp did not think that the Functional Area Assessments contained statistics justifying the need for an ACDB. (The Functional Area Assessments were a series of briefings and meetings that addressed how organizations would comply with Army Regulation (AR)-25-5, Information Management for the Sustaining Base. However, he recommended that the study group speak with Mr. Ron Craven in the Office of Policy and Plans, DISC4.
4. Mr. Haberkamp mentioned a study prepared by American Management Systems (AMS) in the summer of 1986 that identified the level of data redundancy in four or five major Army databases. The study identified the efforts that were required to eliminate data redundancy and improve the data dictionaries.
5. Mr. Haberkamp recommended that the study group call:
 - a. Mr. Steven Cooper of Data Administration Incorporated who might have done some cost/benefit studies on data redundancy.
 - b. Mr. Frank Spielman at the National Institute of Standards and Technology (NIST) who has a bibliography on data administration references.

6. Mr. Haberkamp recommended that the study group review the following documents for more information:
 - a. The December 1988 issue of "Database Programming and Design Magazine".
 - b. The December 1988 issue of "Datamation".
 - c. The February 1986 "Data Administrator's Workshop Proceedings".

AGENCY/ORGANIZATION OF CONTACT: Office of the Director of Information
Systems for Command, Control,
Communications, and Computers (DISC4)

INDIVIDUAL CONTACTED: Maj. C. A. Harris
Analyst
(202) 695-1671

DATE: 31 January 1989

DISCUSSION:

1. The MITRE study group called Maj C. A. Harris upon the recommendation of LtCol Robert DiBona. Maj Harris is with the Office of the Director of Information Systems for Command, Control, Communications, and Computers (DISC4).
2. Maj. Harris is very interested in the topic being pursued by the study group and feels that such a study is long overdue. However, he does not know of any cost-benefit studies in the military sector.
3. Maj. Harris recommended that the study group contact Maj. Michael Napoliello in his organization who is interested in data standardization.

AGENCY/ORGANIZATION OF CONTACT: Office of the Assistant Secretary of the
Army (Financial Management) (ASA(FM))

INDIVIDUAL CONTACTED: Mr. Greg Hoffman
Financial Management System Integration (FMSI)
(317) 543-6647

DATE: 29 December 1988

DISCUSSION:

1. The MITRE study group called Mr. Greg Hoffman upon the recommendation of Mr. James Glymph of the Data Management Directorate (DMD) of the Information Systems Engineering Command (ISEC). Mr. Hoffman works with the Financial Management System Integration (FMSI) group of the Office of the Assistant Secretary of the Army (Financial Management) (ASA(FM)).
2. Mr. Hoffman has been working on the non-standard data elements for the eight financial systems in the Decision Support System (DSS). The DSS is run by the Vice Chief of Staff at the Headquarters of the Department of the Army (HQDA).
3. Presently, Mr. Hoffman is defining and setting-up a thesaurus of elements. Part of this task includes bounding the extent of data non-standardization in quantifiable terms. Mr. Hoffman has no knowledge of any work done in this area. The study will use the staffing conventions included in Army Regulation (AR) 25-9, Army Data Management and Standards Program as a guide.

AGENCY/ORGANIZATION OF CONTACT: Personnel Information Systems Command
(PERSINCOM)

INDIVIDUAL CONTACTED: LtCol Pen Hollist
Program Manager for Personnel Data Administration
(202) 697-1914

DATE: 31 January 1989

DISCUSSION:

1. The MITRE study group called LtCol Pen Hollist upon the recommendation of LtCol William Reyers of Information Systems Command (ISC). LtCol Pen Hollist is the Program Manager for Personnel Data Administration in the Personnel Information Systems Command (PERSINCOM).
2. LtCol Hollist stays current with Information Systems trade journals and publications and knows of no studies on the value or benefits of "good" information.
3. LtCol Hollist discussed the Total Army Personnel Database (TAPDB) and its development. The TAPDB is a microcosm of the former Army Corporate Database (ACDB) concept. It is used to load, edit, and maintain data, and to rectify problems with Army personnel data. The TAPDB contains entries on 2.7 million Army personnel. The reservists have 1.2 million personnel in their database and have voluntarily abandoned their localized system in lieu of a shared Database Management System (DBMS) because of data standardization.
4. LtCol Hollist offered the following definitions of redundancy:
 - a. Multiple copies of the same file.
 - b. The same data known by two different names.
 - c. Different data known by the same name.
5. LtCol Hollist feels that the Personnel Information Command has a different view of the objectives of data standardization than Information Systems Engineering Command (ISEC) does. ISEC is concerned about what the data element is, while the Personnel Information Command is concerned with how the data will be used.
6. LtCol Hollist is still assessing the potential impact of Army Regulation (AR) 25-9, Army Data Management and Standards Program.
7. The TAPDB has necessitated a brute force analysis of nearly 6000 data elements which represent 90 percent of the relevant personnel

data elements. This is a very labor intensive process but is necessary before the design and coding may begin. The approach is to obtain a consensus among the users of the data element definitions. Approximately 4,500 data elements have been reduced to approximately 2,400 consensus elements. LtCol Hollist feels that redundancy has run between 50 and 75 percent not including multiple copies of files in other databases.

8. LtCol Hollist is very interested in supporting a test case for the study group's approach. Further, he feels that the personnel area offers some worthwhile examples. The sample study could examine operational issues (such as assigning Soldier X to Base Y) or policy issues.

AGENCY/ORGANIZATION OF CONTACT: Office of the Secretary of Defense
(OSD)/Office of the Comptroller,
Information Resources

INDIVIDUAL CONTACTED: Ms. Belkis Leong-Hong
Analyst
(202) 693-2874

DATE: 6 February 1989

DISCUSSION:

1. The MITRE study group called Ms. Belkis Leong-Hong upon the recommendation of Dr. Edgar Sibley of George Mason University (GMU). Ms. Leong-Hong works in the Office of the Secretary of Defense (OSD)/Office of the Comptroller, Information Resources.
2. Ms. Leong-Hong's office reviews and makes decisions about proposed service data management projects. She was involved with the Army Corporate Database (ACDB) project. Ms. Leong-Hong feels that the ACDB was insufficiently justified to OSD. She agrees with LtCol William Reyer's attempt to utilize the data economics approach for justifying future Army database projects.

AGENCY/ORGANIZATION OF CONTACT: Air Force/Directorate of Architecture and Technology, Information Architecture Branch (AF/SCTIA)

INDIVIDUAL CONTACTED: Mr. B. Nguyen
Analyst
(202) 695-5906

DATE: 1 February 1989

DISCUSSION:

1. The MITRE study group called Mr. B. Nguyen upon the recommendation of LtCol W. Cato. Mr. Nguyen works with the Air Force/Directorate of Architecture and Technology, Information Architecture Branch (AF/SCTIA).
2. Mr. Nguyen is very familiar with the issues being investigated by the study group. He presently is working toward getting a large data management system under way in the Air Force. This system includes building a data dictionary.
3. Mr. Nguyen recounted a number of cases where cargo shipment and aircraft schedules could not be tracked because of non-standard and unsynchronized data. He feels this situation will persist until standard data elements and shared data environments are adopted.
4. Mr. Nguyen feels the study group will have a difficult time developing costs. In addition, he feels that evaluating the benefits and consequences will be even more difficult. These tend to be intangible and relatively long-term; and people tend not to wait for the quantified results before making a decision.
5. Mr. Nguyen has worked with a financial accounting system which was able to reduce the 62.5 percent redundancy initially resident in the system. However to Mr. Nguyen's knowledge, no cost savings were calculated.

AGENCY/ORGANIZATION OF CONTACT: Merrill-Lynch

INDIVIDUAL CONTACTED: Mr. Daniel Rice
Manager of Information Analysis
(212) 236-3465

DATE: 17 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Daniel Rice after reading an article he wrote entitled, Putting Top-Down and Bottom-Up Analysis Together. Mr. Rice is the Manager of Information Analysis for Merrill-Lynch.
2. Mr. Rice feels that some redundancy is necessary but should be specifically justified. It is his experience that as much as 75 percent redundancy may be unnecessary.
3. Mr. Rice recommends that the study group contact Mr. Eric Vesseley an independent consultant, and Mr. Jose Amasachi of Southern California Edison.

AGENCY/ORGANIZATION OF CONTACT: Information Robotics Group, National
Science Foundation (NSF)

INDIVIDUAL CONTACTED: Mr. Larry Rosenberg
Economist
(202) 357-9592

DATE: 3 February 1989

DISCUSSION:

1. The MITRE study group called Mr. Larry Rosenberg upon the recommendation of Dr. Dan Newlon at National Science Foundation (NSF). Mr. Rosenberg works with the Information Robotics Group at NSF.
2. Mr. Rosenberg does not feel that the problem of quantifying the value of benefits associated with reducing data redundancy has been addressed. He is aware of some studies on information and productivity done by the National Planning Association.

AGENCY/ORGANIZATION OF CONTACT: Information/Database Management Systems
Department, George Mason University

INDIVIDUAL CONTACTED: Dr. Edgar Sibley
Professor
(703) 323-2779

DATE: 24 January 1989

DISCUSSION:

1. The MITRE study group called Dr. Edgar Sibley upon the recommendation of Mr. James Glymph, Information Systems Engineering Command (ISEC). Dr. Sibley is a professor with the Information/Database Management Systems Department of George Mason University. He has served as a consultant to the General Accounting Office (GAO), the United States Army, and the Institute for Defense Analyses (IDA).
2. Dr. Sibley feels that quantifying the benefits of a shared database environment is not likely to be credible. Considering the large price tags associated with today's systems, any cost savings would probably be in the noise and within the band of uncertainty of the estimate.
3. Dr. Sibley recommended that the study group review his book Information Resources and Data Dictionaries and speak with Ms. Belkis Leong-Hong at the Department of Defense Office of the Comptroller.

AGENCY/ORGANIZATION OF CONTACT: Navy Personnel Research and Development Center

INDIVIDUAL CONTACTED: Mr. Joseph Silverman
Information Specialist
(619) 553-8034

DATE: 13 February 1989

DISCUSSION:

1. The MITRE study group called Mr. Joseph Silverman upon the recommendation of Mr. Randy Simpson. Mr. Silverman works for the Naval Personnel Research and Development Center.
2. Mr. Silverman is an experienced Management Information System (MIS) practitioner and is "painfully" aware of the nature of the study group's task. He feels that the quantifiable benefits of management information is an uncharted territory.
3. Mr. Silverman feels that the problem with attributing decision outcomes to particular data elements and information is that no one knows what the true basis of a decision really is.
4. Mr. Silverman reminded the study group that updating data does not ensure correct decisions. Data reflects historical experience but decisions are applied to future outcomes which may respond to unanticipated conditions. Therefore, he feels that the study group should apply their approach to a specific, illustrative example because general solutions deal with too many unsubstantiated assumptions.
5. Mr. Silverman's organization is very concerned with the cost-benefit area and is about to undertake a one to two staff year effort to develop a metric for the Navy to use to assess user satisfaction.

AGENCY/ORGANIZATION OF CONTACT: Private Consultant

INDIVIDUAL CONTACTED: Mr. Randy Simpson
(619) 459-1220

DATE: 12 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Randy Simpson upon the recommendation of Dr. Thomas Varley of the Institute for Defense Analyses (IDA) and Mr. Neil Glassman of the Office for Naval Research (ONR). Mr. Simpson has a background in Economics and Decision Sciences. He is the former Chief of the Decision Analysis/Operations Research Group at ONR. He is presently an independent consultant.
2. Mr. Simpson does not feel that decisions are good or bad. Rather than say a "good" decision was made, it is more accurate to say the "best" decision was made given the information available. Information should be defined as having a defined distribution and confidence level. The attributes of information are less than certain because statistically you are only dealing with a sample. The quality of information may be improved, but usually at a cost. Decision-makers should be careful not to expend more resources improving the quality of the information than they gain in benefits. The payoffs for improving information should be included in the utility function the decision-maker attempts to maximize. Resources may be expended to enhance the payoff associated with the decision by improving the database elements, the database organization, or reducing/eliminating redundancy.
3. The Naval Personnel Research and Development Laboratory is interested in research on personal decision-making (a point of contact is Mr. Joseph Silverman).
4. Mr. Simpson recommends that the study group contact the following people:
 - a. Kenneth Arrow at Stanford University
 - b. Ronald Howard at Stanford University
 - c. Peter Fishburn at Bell Labs
 - d. Peter Farquhan at Carnegie-Mellon University
 - e. Daniel Kenneman at Carnegie-Mellon University

- f. Amos Tuersky at Stanford University
 - g. James Dyer at the University of Texas
 - h. James G. Smith at the Office for Naval Research.
5. Mr. Simpson recommends that the study group review Multiple Criteria Decision-Making by Ralph Keeney.

AGENCY/ORGANIZATION OF CONTACT: National Institute of Standards and
Technology (NIST)

INDIVIDUAL CONTACTED: Mr. Frank Spielman
Chairman, Data Administration Workshop
(301) 975-3257

DATE: 11 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Frank Spielman upon the recommendation of Mr. Bruce Haberkamp. Mr. Spielman works for the National Institute of Standards and Technology (NIST).
2. Mr. Spielman knows of no prior studies that have addressed the topic of this study. He recommends that the study group review two articles in the Data Administrator's Workshop Proceedings. In addition, he will provide the study group with a bibliography that he has compiled about data administration articles.
3. Mr. Spielman recommends that the study group contact Mr. Richard Nauer at Mobil. Mr. Nauer is the leader of the national Data Administration Management Association (DAMA) group.

AGENCY/ORGANIZATION OF CONTACT: The MITRE Corporation

INDIVIDUAL CONTACTED: Dr. Frank Stech
Group Leader
(703) 883-6711

DATE: 16 December 1988

DISCUSSION:

1. The MITRE study group called Dr. Frank Stech upon the recommendation of Dr. William Hutzler of the MITRE Corporation. Dr. Stech is a Group Leader of the MITRE Corporation's Protocol Interoperability Engineering Group. Dr. Stech's primary field is the psychology of decision-making.
2. Dr. Stech feels that the shakiest part of the study group's objectives is the link between bad data and making bad decisions. Typically the process goes in the reverse order; what decisions need to be made, what data are available, how can the available data be used to make that decision. It is reasonable to track the impacts of bad decisions. However, it is difficult to track bad data. The typical decision-makers do not have all the data that is needed nor enough time to sift through all the data. Hence, they do not maximize utility to achieve the best or optimum solution. Instead, decision-makers often adopt the first feasible solution for which they have no contradictory information.
3. Dr. Stech pointed out that it would be easier to model a formalized decision-making process (such as a promotion board) than a new decision-making process. Formalized decision-making processes have rules and are more structured. New decision-making processes are less likely to have the available information and decision rules that are necessary to make choices. The better defined the rules are, the clearer the decision-making process. Clearly defined rules usually require specific data. When the rules are clearly defined, the decision-making process usually does not affect the outcome. Dr. Stech feels that in order for the study group to achieve its goal, it has to have tangible data to track and avoid getting hung up on the decision-making process.
4. Dr. Stech feels that rather than study the soft concept of bad data/decisions, the study group should focus on an actual cost benefit of two alternative data environments.
5. Dr. Stech feels that it is feasible to track the outcome of bad decisions. This should include the opportunity costs associated with missing opportunities due to proceeding down the wrong path.

6. Dr. Stech recommended that the study group look at the following books for more information:
 - a. Rational Choice by Hogarth
 - b. Rational Choice in an Uncertain World by Dawes
 - c. Judgement Under Uncertainty Heuristics and Biases by Tversky
 - d. Acceptable Risk by Fischhoff et al.
7. Dr. Stech recommended that the study group contact the following individuals for more information:
 - a. Robin Hogarth at the University of Chicago Business School
 - b. Amos Tversky at the Stanford University Business School
 - c. Robin Dawes at Carnegie-Mellon University
 - d. Baruch Fischhoff at Carnegie-Mellon University
 - e. Herbert Simon at Carnegie-Mellon University
 - f. Dr. Thomas Varley at the Institute for Defense Analyses (IDA).
8. Dr. Stech recommended that the study group check into the following periodicals for more information:
 - a. "Cognitive Sciences"
 - b. "Behavioral Sciences"
 - c. "Decision Sciences".
9. Dr. Stech recommended that the study group contact the following organizations for more information:
 - a. The Oregon Research Institute
 - b. The Decision Research Group
 - c. Persuptronics
 - d. The Decision Sciences Corporation.

AGENCY/ORGANIZATION OF CONTACT: United States Army Force Development Agency

INDIVIDUAL CONTACTED: Col Gerald Thrash
(202) 694-5055

DATE: 17 January 1989

DISCUSSION:

1. The MITRE study group called Col Gerard Thrash on the recommendation of LtCol Robert DiBona of Information Systems Engineering Command (ISEC). Col Thrash is in charge of the Force Builder Project for the United States Army Force Development Agency.
2. Col Thrash knows that some statistics were prepared for standardizing data elements within the Force Development arena. Col Wayne Byrd, Office of the Director of Information Systems for Command, Control, Communications, and Computers (DISC4) may have done some work in this area.
3. Col Thrash recommends that the study group contact Ms. Connie Leonard within DISC4.

AGENCY/ORGANIZATION OF CONTACT: Office of Naval Research

INDIVIDUAL CONTACTED: Mr. Andre Von Tilburg
Computer Scientist
(703) 696-4303

DATE: 17 January 1989

DISCUSSION:

1. The MITRE study group called Mr. Andre Von Tilburg upon the recommendation of Mr. Neil Glassman. Mr. Von Tilburg is a computer scientist with the Office of Naval Research (ONR).
2. Mr. Von Tilburg told the study group that ONR is completing a five year initiative in distributed tactical decision-making. He recommended that the study group contact the program manager for this effort, Dr. William Vaughan.

AGENCY/ORGANIZATION OF CONTACT: Institute for Defense Analyses (IDA)

INDIVIDUAL CONTACTED: Dr. Thomas Varley
Consultant
(703) 845-2527

DATE: 22 December 1988

DISCUSSION:

1. The MITRE study group called Dr. Thomas Varley upon the recommendation of Dr. Frank Stech of the MITRE Corporation.
2. Dr. Varley did not have any specific comments on the study group's subject. However, he did recommend that the study group contact the following people:
 - a. Marty Talcott - Private consultant
 - b. Neil Glassman - ONR
 - c. Kenneth Arrow - Stanford University
 - d. Robert Wilson - Yale University
 - e. Randy Simpson - La Jolla

AGENCY/ORGANIZATION OF CONTACT: Management Sciences and Information
Systems Department, University of Texas

INDIVIDUAL CONTACTED: Dr. Andrew Whinston
Professor
(512) 471-8879

DATE: 24 January 1989

DISCUSSION:

1. The MITRE study group called Dr. Andrew Whinston upon the recommendation of Dr. James Dyer of the University of Texas.
2. Dr. Whinston has done research in the area of data completeness and its impact on decisions. He has examined how precise and exact a database should be in terms of the decision that is being supported.
3. Dr. Whinston has developed a logistics model that links various levels of precision in information to the optimal routing of supply logistics. Trade-offs are made between the quality of the data and divergence from optimal in logistic routes. Dr. Whinston will send the study group a copy of his paper documenting this model.

AGENCY/ORGANIZATION OF CONTACT: Stanford Research Institute

INDIVIDUAL CONTACTED: Ms. Geri Wong
(415) 859-2133

DATE: 23 January 1989

DISCUSSION:

1. The MITRE study group called Ms. Geri Wong upon the recommendation of Mr. Spiros Coutavas of the MITRE Corporation. Ms. Wong is with the reports section of the Stanford Research Institute (SRI).
2. Ms. Wong did a literature search of SRI documents and identified a number in the area of decision sciences. However, she is unable to provide the study group copies of these because they are proprietary.
3. Ms. Wong recommended that the study group contact the Data Log group at SRI for a listing of available reports.

REFERENCES

- ADPAC Corporation (1989), **Cost Justification Report Guide**.
- Ahituv, N. (December 1980), **A Systematic Approach Toward Assessing the Value of an Information System**, MIS Quarterly.
- AIRMICS (September 1988), **AIRMICS Research Status Report**, U.S. Army Institute for Research in Management, Information, Communications, and Computer Sciences.
- Anderson, A. et al. (February 1988), **Review of 23 Standard Army MISs**, MTR-88W00036, McLean, VA, The MITRE Corporation.
- Army Audit Agency (28 March 1986), **Army Data Element Standardization Program**, U.S. Army Audit Agency.
- Balakrishnan A. G. and Whinston, A. (December 1988), **Information Issues in Model Selection**, National Science Foundation Research Grant, Unpublished Paper.
- Ballou, D. and Tayi, G. (June 1985), **Reconciliation Process for Data Management in Distributed Environments**, MIS Quarterly.
- Barret, S. and Konsynski, B. (1982), **Interorganization Information Sharing Systems**, MIS Quarterly, Special Issue.
- Baumol, William J. (1965), **Economic Theory and Operations Analysis**, Prentice-Hall, Inc.
- Bierman, H. et al. (1977), **Quantitative Analysis for Business Decisions**, Richard D. Irwin, Inc.
- Black, S. and Marchand D. (1982), **Assessing the Value of Information in Organizations: A Challenge for the 1980's**, The Information Society Journal, Vol. 1, No. 3.
- Brewer, G. D. (January 1975), **An Analyst's View of the Uses and Abuses of Modelling for Decision-making**, RAND Corporation, P-5395.
- Burk, C. F. and Horton, F. W. (1988), **Infomap: A Guide to Discovering Corporate Information Resources**, Prentice-Hall, Inc.
- Carlyle, R. E. (1988), **Where Methodology Falls Short**, Datamation.
- Caudle, Sharon L. (Spring 1988), **IRM: A Look Backward and Forward at the Federal Level**, Information Management Review, Vol. 3, No. 4.

REFERENCES (Continued)

- Chandler, J. S. (March 1982), **A Multiple Criteria Approach for Evaluating Information Systems**, MIS Quarterly.
- Chen, Kan (1968), **Decision Analysis for Industry**, Stanford Research Institute, Unpublished.
- Chesler, L. et al. (October 1969), **The Use of Information in Clinical Problem Solving: A Framework for Analysis**, RAND Corporation, P-4207.
- Chick, Morey J. (February 1986), **Information Value and Cost Measures as Management Tools**, (NBSIR 86-3324).
- Chick, Morey J. (Summer 1986), **Policies and Research Needed for Assessing Risk of Automated Information**, Information Management Review, Vol 2, No. 1.
- Crowston, K. and Treacy, M. (October 1986), **Assessing the Impact of Information Technology on Enterprise Level Performance**, Center for Information Systems Research, Massachusetts Institute of Technology, CISR WP-143.
- Davis, C. and Wetherbe, J. (December 1987), **An Analysis of the Impact of Distributed DP on Organizations in 1980s**, MIS Quarterly.
- Dawes, R. M. (1988), **Rational Choice in an Uncertain World**, Harcourt, Brace, and Jovanovich, Inc.
- De, P. and Sen, A. (September 1984), **A New Methodology for Database Requirements Analysis**, MIS Quarterly.
- Dienemanr, P. F. (January 1966), **Estimating Cost Uncertainty Using Monte Carlo Techniques**, RAND Corporation, RM-4854-PR.
- Ginsberg, A. S. (July 1971), **Decision Analysis in Clinical Patient Management with an Application**, RAND Corporation, R-751-RD/NLM.
- Griffiths, Jose-Marie (1982), **The Value of Information and Related Systems, Products and Services**, Annual Review of Information Science and Technology.
- Hamilton, S. and Chervany, N. L. (September 1981), **Evaluating Information Systems Effectiveness-Part I: Comparing Evaluation Approaches**, MIS Quarterly.
- Hamilton, S. and Chervany, N. L. (December 1981), **Evaluating Information Systems Effectiveness-Part II: Comparing Evaluator Viewpoints**, MIS Quarterly.

REFERENCES (Continued)

Hammond, J. S. (June 1971), **Defense Decision-Making: Prudent Versus Excessive Conservatism**, RAND Corporation, R-715-PR.

Herden, Melody A. (September 1987), **Data Standardization Programs - An Overview**, MITRE-MTR-87W150.

Hogarth, R. and Reder, M. (Spring 1988), **Rational Choice: The Contrast Between Economics and Psychology**, University of Chicago Press.

Horton, F. W. and Pruden, J. S. (Spring 1988), **Benefit-Cost Analysis-A Delphi Approach**, Information Management Review, Vol. 3, No. 4.

HQDA (1 March 1986), **AR 25-1: Army Information Management Program**, Headquarters, Department of the Army.

HQDA (29 January 1988), **Information Mission Planning Guidance**, Army Director of Information Systems for Command, Control, Communications, and Computers.

HQDA (July 1983), **TB 18-115-2; Army Information Processing Standards (AIPS), Data Elements & Codes**, Headquarters, Department of the Army, Technical Bulletin.

HQDA (8 July 1988), **AR 25-9: Army Data Management and Standards Program**, Army Director of Information System for Command, Control, Communications, and Computers.

HQDA (October 1987), **Army Data Architecture**, Headquarters, Department of the Army.

HQDA (March 1986), **AR 25-5: Information Management for the Sustaining Base**, Headquarters, Department of the Army.

Ignizio, J. and Gupta J. **Operations Research in Decision-Making**, Crane, Russak & Company, Inc.

Jonscher, C. (December 1983), **Information Resources and Economic Productivity**, Center for Information Systems Research, Massachusetts Institute of Technology, CISR WP-110.

Keen, P. (March 1981), **Value Analysis: Justifying Decision Support Systems**, MIS Quarterly.

REFERENCES (Continued)

Keeney, R. L. and Raiffa, H. **Decisions With Multiple Objectives: Preferences and Value Tradeoffs**, International Institute for Applied Systems Analysis, John Wiley & Sons, Inc.

King, W. and Rodriguez, J. (September 1978), **Evaluating Management Information Systems**, MIS Quarterly.

Kuhn, I. (February 1986), **Centralized Versus Centralized Data Environment**, U.S. Department of Commerce.

Kull, David (July 1987), **Data Dictionaries Can Point the Way**, Computer & Communications Decisions.

Levitan, Karen (1982), **Information Resources Management-IRM**, Annual Review of Information Science and Technology.

Lobley, D. (March-April 1981), **Databases and the Role of the Data Dictionary**, Cost and Management.

MacCrimmon, K. R. (December 1968), **Decision-Making Among Multiple-Attribute Alternatives: A Survey**, RAND Corporation RM-4823-ARPA.

Madnick, S. and Wang, R. (September 1987), **Integrating Disparate Databases for Composite Answers**, Center for Information Systems Research, Massachusetts Institute of Technology, CISR WP-171.

Madnick, S. and Wang, R. (May 1988), **Evolution Towards Strategic Application of Databases Through Composite Information Systems**, Center for Information Systems Research, Massachusetts Institute of Technology, CISR WP-170.

Matlin, G. (September 1979), **What is the Value of Investment in Information Systems?**, MIS Quarterly.

McCall, J. J. (October 1964), **The Economics of Information and Optimal Stopping Rules**, RAND Corporation, P-2987.

Mendenhall, J. H. et al. (30 September 1988), **Basis for STAMIS Integration**, MITRE WP-88W00388.

Nelson, R./Winter, S. (August 1960), **Weather Information and Economic Decisions**, RAND Corporation, RM-2620-NASA.

OMB (June 1988), **OMB Circular A-11: Preparation and Submission of Budget Estimates**, Office of Management and Budget.

REFERENCES (Concluded)

- OMB (December 1985), **OMB Circular A-130: Management of Federal Information Resources**, Office Of Management and Budget.
- Pueschel, L. (February 1988). **Guidelines for an Economic Analysis for the Army Corporate Database**, MITRE, Unpublished report.
- Rensselaer, C. V. (October 1988), **Real World Data Management, Computer Decisions**.
- Rice, D. and Laufer, S. (December 1988), **Putting Top-Down and Bottom-Up Analysis Together**, Database Programming & Design.
- Roberts, F. S. (September 1970), **What if Utility Functions Do Not Exist?**, RAND Corporation, R-528-ARPA.
- Ross, Ronald (1981), **Data Dictionary & Data Administration**, American Management Association.
- Srinivasan, A. (September 1985), **Alternative Measures of System Effectiveness: Associations and Implications**, MIS Quarterly.
- Taylor, Robert (1986), **Value-Added Processes in Information Systems**, Ablex Publishing Corporation.
- U.S. Congress (11 December 1980), **Public Law No. 96-511 The Paperwork Reduction Act of 1980**.
- Vanecek, M. et al. (March 1983), **The Data Dictionary: An Evaluation from the EDP Audit Perspective**, MIS Quarterly.
- Vincent, D. R. (1983), **Information Economics**, Information Executive, Vol. 1.
- Voell, R. (February 1986), **Cost Justifying the Data Administration Function**, National Bureau of Standards NBIR-86-3324.
- Wood, Helen (February 1986), **National & Federal Standards Effort**, National Bureau of Standards, NBIR-86-3324.
- Yang, A. et al. (February 1987), **Cost Benefit Analysis of Intercomponent Data Access Alternatives**, Federal Computer Performance Evaluation and Simulation Center (FEDSIM).

GLOSSARY

Acronym

ACDB	Army Corporate Database
ADP	Automated Data Processing
AF/SCTIA	Air Force Directorate of Architecture and Technology, Information Architecture Branch
AIRMICS	U.S. Army Institute for Research in Management, Information, Communications, and Computer Sciences
AMS	American Management Systems
AR	Army Regulation
ARIST	Annual Review of Information Science and Technology
ARPERCEN	Army Reserve Personnel Center
ASA (FM)	Assistant Secretary of the Army, Financial Management
C4	Command, Control, Communications and Computers
CASE	Computer Aided Software Engineering
C/B	Cost Benefit Ratio
CBA	Cost Benefit Analysis
CER	Cost Estimating Relationship
CIS	Composite Information System
CISR	Center for Information Systems Research
CIVPERCEN	Civilian Personnel Center
COL	Colonel
DA	Data Administrator
DAMA	Data Administration Management Association
DBA	Database Administrative
DBMS	Database Management System
DDN	Defense Data Network
DDP	Distributed Data Processing
DISC4	Directorate of Information Systems for Command, Control, Communications, and Computers
DMD	Data Management Directorate
DOD	Department of Defense
DOS	Department of State
DSS	Decision Support System
EDP	Electronic Data Processing
FAIS	Foreign Affairs Information System
FMSI	Foreign Management System Integration
FSIMC	Federal System Integration and Management Center
FY	Fiscal Year

GLOSSARY (Continued)

Acronym

GAO	General Accounting Office
GMU	George Mason University
GSA	General Services Administration
HP	Hewlett-Packard
HQDA	Headquarters Department of Army
IDA	Institute for Defense Analyses
ILCA	Information Life Cycle Analysis
IMA	Information Mission Area
IRM	Information Resources Management
IS	Information Sharing
ISC	Information Systems Command
ISEC	Information Systems Engineering Command
LTCOL	Lieutenant Colonel
M	Million
MAJ	Major
MILPERCEN	Military Personnel Center
MIS	Management Information System
MIT	Massachusetts Institute of Technology
NAPA	National Association of Public Administrators
NBIR	National Bureau of Standards Information Report
NGB	National Guard Bureau
NIST	National Institute of Standards and Technology
NPRDC	Naval Personnel Research and Development Center
NPV	Net Present Value
NSF	National Science Foundation
OMB	Office of Management and Budget
ONR	Office of Naval Research
OR	Operations Research
OSD	Office of the Secretary of Defense
PEO	Program Executive Officer
PERSINCOM	Personnel Systems Information System Command
PM	Program Manager
R&D	Research and Development
SRI	Stanford Research Institute

GLOSSARY (Continued)

Acronym

STAMIS	Standard Army Management Information System
STAMOD	STAMIS Modernization
TAPDB	Total Army Personnel Database
USA	United States Army
UT	University of Texas

Terms

Alias	A name attributed to an object other than its primary or preferred name.
Army Corporate Database (ACDB)	A concept originated in September 1983 by the Army Vice Chief of Staff designed to make Army data a resource available to all who need it. A corporate database is a database characterized by a high degree of integration in data structuring and a high degree of centralization in usage and administration. The goal of the ACDB was to "... stop operating a multitude of private stove-pipe databases which are accessible only to the proponent."
Data Architecture	A framework for organizing data to support the way an entity does or wants to do business. A catalogue or notation representing an entity occurrence complex used to define and manage a configuration of subject areas, functions, processes, and/or activities, and to control any changes to them.
Data Administrator	A person or team responsible for planning, coordinating, and protecting the data resources of the entity, using the facilities and concepts of DDSs and DBMSs in particular.
Database	A collection of interrelated, largely unique data items or records, in one or more computer files, which may be processed by many different applications programs. A database is created and maintained by a DBMS.

GLOSSARY (Continued)

Terms

Database Management System	A computerized system consisting of numerous components, which have as their collective purposes the implementation, management, and protection of databases.
Data Dictionary	A repository for definitions and related information for data resources of an organization.
Data Dictionary System (DDS)	An automated system for the management of the data dictionary.
Data Element	The logical definition of a unit of information within the entity's data resources, apart from its actual use within any given program, file database report, screen etc.
Data Element Dictionary	A repository for definitions and related information for the data resources of an entity.
Data Element Supplier	The entity which captures and enters data into an entity's database.
Data Independence	A characteristic of database systems arising from the segregation of data structure definition from data access (or application programs).
Data Model	A conceptual representation of data, how they are used, and how they are interrelated.
Data Redundancy	A characteristic of traditional master file systems, in which duplicate data are often carried by two or more files; also, the appearance of a given data element in two or more stored files or databases.
Data Resource	The information on the basis of which activities are planned, coordinated, executed, and evaluated.
Data Sharing	The ability of several users to access common data. This ability reduces unwanted data redundancy and inconsistencies.

GLOSSARY (Continued)

Terms

Data Standard	A collection of attributes about a fact or assumption which was subjected to prescribed structuring criteria and approved by an appropriate official for use.
Decentralized Data Bases	Databases which are dispersed to support local processing at individual locations without the controlling influence of a centrally managed coordination effort.
Decentralized Data Processing	A data processing environment where the computing facilities and possibly the responsibility are dispersed in an uncoordinated manner.
Decision Theory	Mathematical modeling of maximizing behavior under uncertainty where the maximizing player faces a nature or environment about which he knows nothing (unlike game theory, which models player conflicts where behavior is more predictable).
Delphi	A process involving the pooling of expert opinions based on a rigorous methodology that channels varying views to convergence.
Distributed Data Base	Databases dispersed to support local processing at individual locations with the controlling influence of a centrally managed coordination effort to ensure that data is properly interrelated.
Distributed Data Processing	Using a collection of individual machines, usually in a network, to satisfy the processing requirements of corporate data resource management.
Effectiveness	Measurement of an information system performance that focuses upon the impact of information output on user performance.
Efficiency	Measurement of an information system's performance that narrowly focuses on the data processing functions (store, retrieval, and transfer) of the information life cycle.

GLOSSARY (Continued)

Terms

Game Theory	Mathematical modeling of maximizing behavior under conflict and risk, where the behavior of other players has a major element of predictability (i.e., others act to reduce the benefits of the first player).
Information Resources Management (IRM)	An approach, fostered by the Paperwork Reduction Act of 1980, to assure that information produced from information resources has maximum value and is produced at the most efficient cost. IRM applies appropriate management philosophy, methodology and techniques to decisions about data, information, and other information resources (equipment, software, personnel, etc.).
Inter-Organizational Information Sharing System	A general term referring to a system that involves resources shared between two or more organizations.
Integration	The elimination of partially overlapping data management schemes, and substitution of a central data structure in their place
Integrity	Data accuracy and consistency.
Law of Diminishing Marginal Return (Utility)	As used in this report, it describes the process by which each successive unit of acquired information yields less extra benefit or utility than the previous one, as want for more information comes closer to fulfillment, and the cost of obtaining more information increases at an increasing rate.
Management Information System	A computer system oriented toward producing decision data, that is, data useful to centralized administration, planning, policy formulation, and so forth. An MIS implies summarization and interpretation of large amounts of "raw" data.
Metadata	Data that describe and define other data. For example, a data dictionary contains a collection of metadata.

GLOSSARY (Concluded)

Terms

Multiple-attribute decisions	Multiple properties or dimensions by which a decision's consequences are measured or described - e.g., dollars, time units, percentage of objectives met, etc.
Network	A collection of communicating machines.
Operations Research	Application of scientific methods and techniques to decision-making problems where there are two or more alternative courses of action leading to different and often unknown results.
Risk	Refers to situations in which the outcome is not certain, but where the probabilities of the alternatives are known, or can at least be estimated.
Satisfice	A term coined by Herbert Simon that describes the "real" behavior of decision-makers in maximization problems, specifically rejecting the notion that they expend the time and energy to sift through all information and options to maximize. Instead, decision-makers "satisfice", or look for solutions that are acceptable, reasonable, or satisfactory and which require only a moderate review of options.
Stove-Pipe System	An Army originated term which means that a given Army functional component develops a database and submits required data to higher or lower levels. A stove-pipe system is a database system containing data applicable to limited functions areas.
Uncertainty	Refers to situations where outcomes are not certain and cannot even be predicted in probabilistic terms.
Utility	Used to describe the consequences of an action or decision--a non-dimensional measure of the decision-maker's want-satisfying, psychological response to an outcome.

DISTRIBUTION LIST

INTERNAL

A-10 G. J. MacDonald
D-14 R. M. Harris
J. S. Quilty
W-110 J. Dominitz
J. C. Slaybaugh
W-111 A. H. Mossler
J. H. Wood
W-113 C. A. Castellana (3)
T. J. Coonce (3)
S. D. Coutavas (3)
W. P. Hutzler
J. S. Lovelace
C. V. Moran
W-114 A. K. Anderson
R. J. Dube, Jr.
R. Goldgraben
P. T. Hedeman
J. Hustad
M. G. Hutchinson
S. E. McLeod
C. G. Miller
R. J. Nieporent
R. T. Nixon
T. N. Nguyen
E. M. Ozarko
R. A. Ringdahl (3)
W. B. Stevens (2)

Records Resources (2)
Project Files (1)

EXTERNAL

Commander
United States Army Infor-
mation Systems Command
Fort Huachuca, AZ 85613
COL E. V. Freeman, ASPL
LTC W. Reyers, ASPL-AD (2)
R. J. Furlong, ASPL-PS
M. Gentry, ASTD-A
B. Kappes, ASPL
L. J. Mabius, ASTD
E. A. Singer, ASPL-PT
H. Slater, ASPL-PT
R. Walker, ASPL-PT (2)
I. B. Zundel, ASPL-P
Commander
United States Army Infor-
mation System Engineering
Command Fort Huachuca,
AZ 85613
R. Boggie, ASQB-SED
J. Hunter, ASQB-TDA
J. Rudigier, ASQB-SID
W. Wachsman, ASQB-SAO